

Requested Patent: JP6319563A

Title:

HEPATITIS C VIRAL GENE, OLIGONUCLEOTIDE AND METHOD FOR  
DETERMINING HEPATITIS C VIRAL GENOTYPE ;

Abstracted Patent: JP6319563 ;

Publication Date: 1994-11-22 ;

Inventor(s): OKAMOTO HIROAKI; others: 01 ;

Applicant(s): IMUNO JAPAN:KK ;

Application Number: JP19930147133 19930513 ;

Priority Number(s): ;

IPC Classification:

C12N15/51 ; C07K7/08 ; C07K13/00 ; C12P21/02 ; C12Q1/68 ; G01N33/576 ;

Equivalents:

ABSTRACT:

PURPOSE:To obtain a gene of a newly elucidated hepatitis C virus (HCV), a specific oligonucleotide, a method for detection thereof and a method for determining the genotype of the HCV.

CONSTITUTION:The polynucleotide has a base sequence described in sequence Nos. 1 to 5. The oligonucleotide is described in sequence No 6. The methods for detecting an HCV gene and determining the genotype use the polynucleotide and oligonucleotide as a primer, a probe, etc. Furthermore, the polypeptide is described in sequence Nos. 15 to 19. Thereby, the HCV gene of the newly found genotype can be detected and the HCV genotype Can simultaneously be determined over a wide range.

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平6-319563

(43) 公開日 平成6年(1994)11月22日

(51) Int.Cl. <sup>5</sup>	識別記号	庁内整理番号	F I	技術表示箇所
C 1 2 N 15/51	Z N A			
C 0 7 K 7/08		8318-4H		
13/00		8318-4H		
C 1 2 P 21/02	C	8214-4B		
		9050-4B		
			C 1 2 N 15/ 00	A

審査請求 未請求 請求項の数22 書面 (全 41 頁) 最終頁に続く

(21) 出願番号 特願平5-147133

(22) 出願日 平成5年(1993)5月13日

(71) 出願人 391039391

株式会社イムノ・ジャパン

東京都杉並区荻窪4丁目28番14-701号

(72) 発明者 岡本 宏明

栃木県下都賀郡石橋町石橋1560-25

(72) 発明者 中村 徹雄

東京都杉並区荻窪4丁目28番14-701号

(74) 代理人 弁理士 中島 敏

(54) 【発明の名称】 C型肝炎ウイルス遺伝子、オリゴヌクレオチド、並びにC型 肝炎ウイルス遺伝子型判定方法

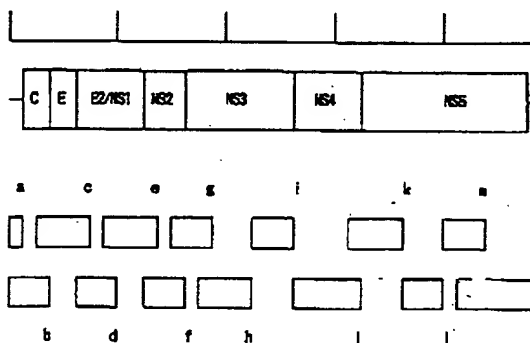
(57) 【要約】

【目的】 新規に解明されたC型肝炎ウイルス (HCV) の遺伝子、特異オリゴヌクレオチド、これら検出する方法、HCVの遺伝子型を判定する方法等を提供することを目的とする。

【構成】 配列番号1ないし5記載の塩基配列を有するポリヌクレオチド、配列番号6記載のオリゴヌクレオチド、これらをプライマー、プローブ等として使用するHCV遺伝子の検出法、遺伝子型判定法の発明、ならびに配列番号15ないし19記載のポリペプチドの発明である。

【効果】 新たに発見された遺伝子型のHCV遺伝子を検出するとともに、広い範囲に涉ってHCV遺伝子型を判定することができる。

塩基配列決定に利用したHCV領域



a (nt1-160) ; b (nt63-847) ; c (nt732-1806) ; d (nt1300-1857) ; e (nt1788-2560)  
f (nt241-3018) ; g (nt2973-3383) ; h (nt3723-4737) ; i (nt4833-5050)  
j (nt5003-6170) ; k (nt6129-7086) ; l (nt7023-7833) ; m (nt7792-8530)  
n (nt8253-9440)

左端に5'端、右に3'端が示されている。

塩基配列は5'端からの塩基数 (nt) で示した。

## 【特許請求の範囲】

【請求項1】配列番号1記載の塩基配列、またはこれと相補的な塩基配列を有するC型肝炎ウイルス遺伝子cDNA・HC-G9。

【請求項2】配列番号2記載の塩基配列、またはこれと相補的な塩基配列を有するポリヌクレオチドYS117・5'。

【請求項3】配列番号3記載の塩基配列、またはこれと相補的な塩基配列を有するポリヌクレオチドYS117・3'。

【請求項4】配列番号4記載の塩基配列、またはこれと相補的な塩基配列を有するポリヌクレオチドSR037・5'。

【請求項5】配列番号5記載の塩基配列、またはこれと相補的な塩基配列を有するポリヌクレオチドSR037・3'。

【請求項6】配列番号1～5記載の塩基配列の一部、またはこれと相補的な塩基配列の一部を構成するポリヌクレオチドまたはオリゴヌクレオチド。

【請求項7】請求項第6項記載のオリゴヌクレオチドからなるプライマーまたはプライマーペア。

【請求項8】配列番号6記載の塩基配列を有するオリゴヌクレオチド#321。

【請求項9】配列番号6記載の塩基配列を有するオリゴヌクレオチドプライマー。

【請求項10】配列番号6記載の塩基配列を有するオリゴヌクレオチドプライマーと配列番号7ないし9記載の塩基配列を有するオリゴヌクレオチドプライマーの1種以上からなるプライマーペア。

【請求項11】配列番号6記載の塩基配列を有するオリゴヌクレオチドプライマーと配列番号10ないし14記載の塩基配列を有するオリゴヌクレオチドプライマーの1種以上からなるプライマーペア。

【請求項12】請求項第7項ないし第11項記載のオリゴヌクレオチドプライマーまたはオリゴヌクレオチドプライマーペアを使用するC型肝炎ウイルス遺伝子型判定法。

【請求項13】請求項第7項ないし第10項記載のオリゴヌクレオチドプライマーペアを使用するC型肝炎ウイルスの遺伝子検出法。

【請求項14】ポリメラーゼチェーンリアクション法によりcDNAを増幅することを特徴とする請求項第12項記載のC型肝炎ウイルス遺伝子型判定法または請求項第13項記載のC型肝炎ウイルス遺伝子検出法。

【請求項15】請求項第11項または第12項記載のオリゴヌクレオチドプライマーを複数組使用し、複数回の増幅を行う請求項第13項記載のC型肝炎ウイルス遺伝子検出法または第14項記載のC型肝炎ウイルス遺伝子型判定法。

【請求項16】請求項第6項ないし第9項記載の標識プ

ロープ。

【請求項17】請求項第16項記載の標識プローブを使用したC型肝炎ウイルス遺伝子型判定法。

【請求項18】配列番号15記載のポリペプチドHC-G9 Protein、ならびにその部分ペプチド。

【請求項19】配列番号16記載のポリペプチドYS117・5' Peptide、ならびにその部分ペプチド。

【請求項20】配列番号17記載のポリペプチドYS117・3' Peptide、ならびにその部分ペプチド。

【請求項21】配列番号18記載のポリペプチドSR037・5' Peptide、ならびにその部分ペプチド。

【請求項22】配列番号19記載のポリペプチドSR037・3' Peptide、ならびにその部分ペプチド。

## 【発明の詳細な説明】

## 【0001】

【産業上の利用分野】本発明は、C型肝炎ウイルス（以下「HCV」と略記する）のcDNA遺伝子、これを構成する特異オリゴヌクレオチド、これに由来する蛋白質、ペプチド、ならびにこれらを用いたHCV遺伝子型判定法に関する。

## 【0002】

【従来の技術】1988年にHCV遺伝子の一部が解明され発表されて以来、HCV診断への応用可能な数多くの技術が開発され実用化されてきた。これまで、HCV感染によって患者血液中に現われるHCV抗体を検出する抗体検査法、ならびに体内に存在するHCV遺伝子を検出する方法が開発され広く用いられてきたが、さらにHCVの遺伝子型を判定する方法が研究、開発され、本発明者らもこれに深く関与してきた。これらの診断技術に於ける現時点の最重要課題は、高い感度と各遺伝子型に対する高い特異性を実現することである。この技術課題を解決する為には、HCVの各遺伝子型に特異的な遺伝子配列あるいは特異抗原の特定とそれに基づく診断技術の確立が急務である。実際に、いくつかのHCV株については遺伝子の全配列が解明されており、さらに他の株については遺伝子配列の一部が解明され、HCV特異遺伝子の配列の特定あるいはHCV特異アミノ酸配列の特定に利用された。その結果、従来の検査法に比べ、これらの情報に基づいて開発された最近の診断法は高い特異性と感度を有するようになり、これにもとづいて、適切な治療方針を採用できるようになりはじめた。しかし、他方、これらの検査法を用いた場合でも、なお捕捉できないHCV疾患例があることも判明しており、より高い特異性と感度を有する診断法の開発が望まれている。

## 【0003】

3

【発明が解決しようとする課題】HCVはその遺伝子配列が初めて解明されてからまだ時間が浅く、ウイルス本体は未だ確認されていない。また、全遺伝子配列が解明されたHCV株もまだ少数であり、HCVの遺伝特性を完全に解明したとは言えないのが実態である。したがって、現在までに発表された遺伝子配列が全てのHCVに共通の情報を提供しているものか否かは明らかでない。完全なHCV診断法を完成させるためには、現在の診断法では捕捉できないHCVの遺伝子特性を解明し、その情報を反映させた診断法の構築が不可欠である。本発明の目的は、今日までのHCV検査法では十分に特徴付けられないHCV株の遺伝子配列を明かにするとともに、その遺伝子特性を解明することにより、正確かつ広範な適用範囲をもつ遺伝子型判定法と、これに用いるオリゴヌクレオチド等を提供することである。

【0004】

【課題を解決するための手段】本発明者らは、今日までの各種HCVの検査では十分に特徴付けられないHCV株の遺伝子本体の解明を目的として、鋭意研究を進めたが、その結果、HCV-RNA陽性でありながら、従来の遺伝子型判定法では型判定できなかったヒト検体からRNAを単離し、これを用いて未知のHCV株の全遺伝子配列を決定した。さらに本発明者らは、この新規の遺伝子配列と従来報告された各遺伝子型の公知のHCV遺伝子配列とを比較した結果、本発明のHCV株が公知の遺伝子型のいずれに相当するものでなく、全く別の新しい遺伝子型であることを解明した（発明者らは暫定的にこの遺伝子型を1c型と命名した）。これに基づいて、本発明者らは、1c型遺伝子型に特異的で他の遺伝子型には存在しない遺伝子配列を特定した。この配列を有するオリゴヌクレオチドをプライマーあるいはプローブとして使用することにより1c型の遺伝子型判定が可能になった。さらに公知の遺伝子型を含めた全ての遺伝子型に共通な遺伝子配列を特定するとともに、この配列を有するオリゴヌクレオチドと1c型に特異的な配列を有する本発明のオリゴヌクレオチドをプライマーとして利用することにより一度の検査で遺伝子型の判定が実現できることを見出し、本発明を完成した。本発明のポリヌクレオチドおよびオリゴヌクレオチドからなるプライマーならびにプローブは、HCV各遺伝子型に共通する遺伝子あるいは型特異的な遺伝子配列に対して特異的に結合することにより作用を発揮するものである。ポリヌクレオチドならびにオリゴヌクレオチドの結合には配列上若干の差異があっても影響を受けないことは周知のことであるから、本発明のポリヌクレオチドまたはオリゴヌクレオチドに対して若干の置換を有するポリヌクレオチドならびにオリゴヌクレオチドも当然本発明の範囲に包含される。

【0005】すなわち、本発明は公知のHCVとは異なる新しい遺伝子型である1c型を有するHCVたるHCV

4

-G9のcDNA遺伝子の発明であり、またその特異的な遺伝子配列の一部を構成する、あるいはこれに相補的な塩基配列を有するポリヌクレオチドまたはオリゴヌクレオチドに関する発明であり、具体的には配列番号1ないし5記載の塩基配列、またはこれに相補的な塩基配列の全部あるいは一部を有するポリヌクレオチドあるいはオリゴヌクレオチドの発明である。本発明は、上記オリゴヌクレオチドからなるプライマー、プローブ、あるいは標識されたプローブの発明であり、また配列番号6記載の塩基配列の全部あるいは一部を有するプライマーあるいはプローブに関する発明である。本発明は、上記プライマーと、配列番号7ないし14記載の遺伝子配列を有するオリゴヌクレオチドからなるプライマーとを組み合わせることで利用し、HCV遺伝子の検出、遺伝子型の判定を行うことができる混合プライマーに関する発明である。また、本発明は上記プライマーあるいはプローブを単独、あるいは同時に使用することによってHCVの遺伝子を検出する方法、または遺伝子型を判定する方法の発明である。

【0006】本発明者らは、実施例1に示すように、従来の遺伝子型判定法では型判定できなかったHCV抗体陽性の複数のヒト検体より所定の方法でRNAを抽出し、HC-G9については全遺伝子配列を特定し、該HCVをHC-G9と命名し、残りの2検体についてはその一部の配列を特定し、該検体をS117、SR037と命名した。その塩基配列は配列番号1ないし5記載のとおりである。HC-G9の全遺伝子配列は3'端側に見られたTストレッチ部分を除いて9440個の塩基から成り、5'端に341塩基からなる非翻訳領域を、続いて9033塩基からなり3011アミノ酸をコードする領域が、更にこれに続く3'端に66塩基からなる非翻訳領域より構成されることが判明した。

【0007】本発明者は、実施例2に示すように本発明の対象である上記各株の塩基配列と公知の4つの遺伝子型に属する14株の塩基配列とを比較した、その結果、上記株の塩基配列と既知の遺伝子型配列との間には20%以上の非相同性があり、本発明の遺伝子およびポリヌクレオチドは、いずれの公知遺伝子型にも分類できない新規のものであることが見出された。さらに、本発明者らは、本発明の遺伝子およびポリヌクレオチドについて、一部の遺伝子配列のみが公知であるHCV株に対しても遺伝子配列の比較を行った。その結果、本発明にかかる遺伝子およびポリヌクレオチドは、上記一部配列のみ判明しているHCVとも別型である独立した遺伝子型として分類されることを見出した。本発明者らは本発明にかかる新規のHCVの遺伝子型を暫定的に1c型と命名した。

【0008】本発明者らは、実施例2に示すように1c型HCV株の遺伝子型判定に最適な遺伝子領域としてコア領域を特定した。この領域から1c型特異的な配列を有

5

する配列番号6記載のオリゴヌクレオチドを得た。

【0009】本発明者らは、HCVの遺伝子検出および遺伝子型判定に用いるべき領域が、1c型遺伝子型においても他の遺伝子型判定の場合と同様にコア領域にあることに注目し、配列番号6記載の本発明のポリヌクレオチドと併用することができるポリヌクレオチドを公知のプライマーのなかから検索した。その結果、配列番号7記載の公知のオリゴヌクレオチドの配列は1c型にもよく保存されており、これがHCVの遺伝子検出および遺伝子型判定に於ける共通プライマーとして利用可能なことが見出した。配列番号6記載の1c型特異的オリゴヌクレオチドプライマーを公知の共通プライマーたる配列番号7記載のオリゴヌクレオチドプライマーとを組合せて使用することによって、1c型の遺伝子を特異的に増幅できることを見出した。

【0010】本発明では遺伝子の増幅方法としてポリメラーゼ、チエイン、リアクション法（PCR法）を好適に利用することができる。その際、プライマーペアとしては、配列番号6記載のオリゴヌクレオチドと配列番号9記載のプライマーペアの組合せも好適である。

【0011】PCR法に於いては、第一段階として各遺伝子型に共通な領域を増幅し、第二段階として型特異的な増幅を行うことにより、より高感度にHCV遺伝子の検出と遺伝子型の判定がなされる。この様態のPCR法に使用する好適なオリゴヌクレオチドペアとしては、第一段階用として配列番号7記載のオリゴヌクレオチドと配列番号8記載のオリゴヌクレオチド、第二段階用として配列番号6記載のオリゴヌクレオチドと配列番号9記載のオリゴヌクレオチドを例示することができる。

【0012】また、本発明の方法によれば本発明の新規オリゴヌクレオチド（配列番号6）を配列番号7記載ないし配列番号14の各既知遺伝子型特異的オリゴヌクレオチドプライマーまたは、共通オリゴヌクレオチドプライマーとを同時に使用し、PCR法あるいは公知の2段階PCR法使用して、一度の操作によりHCV遺伝子を検出し、また遺伝子型を判定することができる。

【0013】また、本発明者らは公知の宿主に組み込んで発現させ、また常法により化学合成して、本発明のポリペプチドを得た。

【0014】本発明のポリペプチドにおけるアミノ酸配列は、公知の遺伝子型HCVにおけるポリペプチドとコア領域において高い相同性を有するが、エンベロープ（E1）およびE2/NS1ではその相同性が低く、型特異的であることを示している。E1、E2/NS1はウイルス粒子表面に存在する蛋白質と考えられ、この部分に対する抗体を有する症例も少なからず見出される。したがって、本発明のポリペプチドおよびその部分オリゴペプチドは1c型特異的に抗エンベロープ抗体の検出系の作成や、HCVワクチンに使用することができる。またNS2～NS5領域はプロテアーゼ等の非

6

構造蛋白質をコードしていると考えられ、本発明のポリペプチドおよびその部分オリゴペプチドは非特異的な抗体検出系やHCV増殖阻害剤の開発に用いることができる。

【0015】

【作用】本発明は、新たに見いだされた遺伝子型である1c型のHCVを含め、従来よりも広汎な遺伝子型のHCVを高感度に検出し、またそのHCVの各遺伝子型と同時に特異的に判定することができる。本発明の遺伝子、ポリヌクレオチド、オリゴヌクレオチドプライマー、蛋白質、ペプチドはC.V.遺伝子の検出ならびに遺伝子型判定に供することができる。

【0016】

【実施例】以下、本発明の実施例についての述べるが、もとより本発明がこれらの実施例に限定されるものではない。

【0017】実施例1

従来の遺伝子型に分類されない複数の新規HCV株を見出し、その全塩基配列および一部の塩基配列を次のようにして決定した。

【0018】（1）RNAの抽出

市販のHCV抗体検査薬および本発明者らによって開発され特許出願中の抗体検査法（特願平2-153401）ならびに、本発明者らによって別途特許出願されたオリゴヌクレオチドプライマーを用いたHCV検出法（特開平5-23200）によりHCV感染が確認されているが、本発明者により別途特許出願されているHCV遺伝子型判定に関する方法（特願平3-307296、4-093960）ではその遺伝子型判定ができなかった肝炎患者由来の血液検体（HC-G9、YS117、およびSR037）から次のようにしてRNAを抽出した。血清50μlに適当量のトリス緩衝液（10mM、pH8.0）を加え、90×103rpmにて15分間の遠心分離を行った。得られたペレットに200mMのNaCl、10mMEDTA、2%（重量/容積）のドデシル硫酸ナトリウム（SDS）と1mg/mlのプロテナーゼKを含むトリス緩衝液（50mM、pH8.0）を加え、60℃で1時間加温し、フェノール/クロロホルムで抽出を行った後、エタノール沈澱を行いRNAを得た。

【0019】（2）cDNAの作製

各検体より得たRNAを70℃で1時間加温した後、これを急冷し鋳型RNAとした。この鋳型RNAサンプルに100ユニットの逆転写酵素（Superscript；GIBCO，BRL）およびオリゴヌクレオチドプライマー20pmolを加え、42℃、1時間反応させてcDNAを得た。

【0020】（3）cDNAのポリメラーゼチエインリアクション（PCR）による増幅

上記の操作により得られた単鎖cDNAについて、図1

7

8

に示す領域別に、表1に示す各領域毎に設定したセンス側オリゴヌクレオチドプライマーならびにアンチセンス側オリゴヌクレオチドプライマーからなるプライマーペアを用いて増幅を行った。増幅は、DNAサーマルサイクラー (Perkin-Elmer, Cetus) に Gene Amp DNA増幅試薬キット (Perkin-Elmer・Cetus) を用いて Saiki 氏の方法\*

\*法 [Science, Vol. 239, p487-491 (1988)] に従って35回の増幅サイクルからなるPCR法にて実施した。

[0021]

[図1]

[0022]

[表1]

配列番号: 1

配列の長さ: 9487

配列の型: 核酸

鎖の数: 一本鎖

トポロジー: 直線状

配列の種類: cDNA to genomic RNA (HC-G9 cDNA)

```

GCCAGCCCCC TGATGGGGGC GACACTCCGC CATGAATCAC TCCCCTGTGA GGAACACTG 60
TCTTCACGCA GAAAGCGTCT AGCCATGGCG TTAGTATGAG TGTCGTGCAG CCTCCAGGAC 120
CCCCCTCCC GGGAGAGCCA TAGTGGTCTG CGGAACCGGT GAGTACACCG GAATTGCCAG 180
GACGACCGGG TCCTTTCTTG GATTAACCCG CTCAATGCCT GGAGATTTGG GCGTGCCCCC 240
GCAAGACTGC TAGCCGAGTA GTGTTGGGTC GCGAAAGGCC TTGTGGTACT GCCTGATAGG 300
GTGCTTGCGA GTGCCCCGGG AGGTCTCGTA GACCGTGCAC CATGAGCAGC AATCCTAAAC 360
CTCAAAGAAA AACCAAACGT AACACCAACC GCCGCCACA GGACGTTAAG TTCCCGGGTG 420
GCGGCCAGAT CGTTGGCGGA GTTTACTTGT TGCCGCGCAG GGGCCCCAGA GTGGGTGTGC 480
GCGCGACGAG GAAGACTTCC GAGCGGTCCG AACCTCGCGG GAGGCGTCAG CCTATTCCA 540
AGGCCCGCCG ACCCGAGGGA AGGTCTTGGG CGCAGCCCGG GTACCCCTGG CCCCTCTATG 600
GCAACGAGGG CTGTGGGTGG GCGGGATGGC TCCTGTCCCC CCGCGGCTCT CGGCCTAGTT 660
GGGGCCCTTC TGACCCCCGG CGGAGGTAC GCAATTTGGG TAAGGTCATC GATACCCTCA 720
CGTGTGGCIT CGCCGACCTC ATGGGGTACA TCCCCTCGT CCGCGCTCCT CTAGGGGGCG 780
CTGCCAGAGC TCTGGACAT GGTGTTAGAG TCCTGGAAGA CCGCGTGAAT TACGCAACAG 840
GGAACCTCCC CGGTTGCTCT TTTTCTATCT TCTTGCTCGC TCTTCTATCC TGCCTGACAG 900
TCCCTGCTTC GGCCGTCGGA GTGCGCAACT CTTCGGGGGT GTACCATGTC ACCAATGATT 960
GCCCCAATGC GTCCGTTGTG TACGAGACGG AGAACCTGAT CATGCATCTG CCCGGGTGTG 1020
TGCCCTACGT ACGCGAGGGC AACGCCTCGA GGTGTTGGGT CTCCCTTAGT CCCACCGTAG 1080
CCGCCAGGGA TTCGCGCGTC CCCGTCACTG AGGTTCCGGC TCGTGTCCAG TCGATTGTCTG 1140
GGGCCGCTGC GTTCTGTTCT GCTATGTATG TAGGGGACCT ATGCGGCTCC ATCTTCCTTG 1200
TTGGCCAGAT CTTACCTTC TCTCCAGGC ACCATTGGAC GACGCAAGAC TGCAATTGCT 1260
CCATCTACCC AGGCCATGTG ACAGGTCATC GAATGGCTTG GGACATGATG ATGAATTGGT 1320

```

[0023] (4) cDNAライブラリーの構築による HC-G9、YS117、およびSR037の塩基配列の決定 PCRにて増幅した各検体由来の各領域遺伝子をT4ポ

リヌクレオチドカイネース (New England Biolabs)、T4 DNAポリメラーゼ (Takara Biochemicals) で処理後、M13 ファージベクターに挿入し、クローン化した。塩基配列

9

決定はdideoxy chain termination法にて、Sequencenase sequencekit ver 2.0 (United States Biochemicals)あるいはAutoRead Sequencing kit (Pharmacia)を用いて行った。各検体について、各領域3クローンを得、それぞれについて配列を決定し、各クローンに共通する塩基、あるいは一致率の高い塩基を採用して配列を決定した。HC-G9については全領域を、YS117ならびにSR037については5'端より63番目の塩基から1867番目の塩基までの配列と8259番目から9440番目までの配列について決定した。配列表1にHC-G9、配列表2にYS117、の5'側、配列表3にYS117、の3'側、配列表4にSR037の5'側、配列表5にSR037の塩基配列の3'側の各塩基配列を示す。表2にHC-G9、YS117、SR037の配列間の相同性を示す。その結果、これら3検体の塩基配列の相同性は95%あり、上

10

記の3検体のHCVは同一の遺伝子型に分類される。図2にHC-G9とこれまでに全塩基配列が解明されている14のHCV株との塩基配列の相同性を示す。その結果、HC-G9はいずれの既知の遺伝子型株の塩基配列とも20%以上の非相同性をしめすことが判明した。このことからHC-G9を含む3検体由来のHCVは従来の遺伝子型とは異なる新しい遺伝子型に分類されることを確認した。遺伝子配列に基づく系統分類の推定より、この新たに見いだされた遺伝子型は大きく1型と呼ばれる分類に属し、その分類にはすでにI型、II型が存在し、これらは1aならびに1bと呼称されていることから、これに倣って1c型と暫定的に呼ぶこととした。

【0024】

【図2】

【0025】

【表2】

11

12

CACCCACTGG CGCCTTAGTG GTGGCACAGC TACTCCGGAT CCCACAAGCT ATCGTGGATA 1380  
 TGATAGCTGG TGCCCACTGG GGTGTCTAG CGGGCCTGGC ATACTACTCC ATGGTGGGA 1440  
 ACTGGGCTAA GGTGTGGTC GTGCTGCTGC TCTTCGCTGG CATTGACGCA GAGACCCGGG 1500  
 TCACAGGGGG GGGCGCTGGC CACACCGCGT TCGGGTTTGC TAGCTTCCTC GCCCAAGGCG 1560  
 CTAAGCAAAA GATCCAGCTC ATAAATACCA ACGGCAGCTG GCACATCAAC AGAACTGCCT 1620  
 TGAAGTGTA TGAAGCTTG GATACTGGCT GGCTAGCAGG GCTGCTCTAC TACCACAAGT 1680  
 TCAACTCCTC AGGGTGTCCC GAGAGGATGG CTAGTTGCCA ACCTCTTACC GCCTTCGACC 1740  
 AAGGGTGGGG ACCCATCACT CACGAGGGGA ATGCTAGTGA TGACCAGCGG CCATATTGTT 1800  
 GGCACATATC CCTACGCCCG TGTGGCATTG TGCCAGCGAA AAAGGTTTGC GGGCCTGTAT 1860  
 ACTGTTTCAC ACCCAGCCCC GTGGTAGTGG GGACGACGGA CAGAGCCGGC GTTCCTACCT 1920  
 ACAGATGGGG TGCCAATGAG ACGGATGTAC TGCTCTCAA CAACTCTAGG CCGCCAATGG 1980  
 GGAATTGGTT TGGGTGTACG TGGATGAATT CTAGTGGCTT CACCAAGACG TCGGGGGCTC 2040  
 CGGCCTGCAA CATCGGCGGG AGCGGGAACA ATACCCTGCT GTGCCAACA GATTGCTTCC 2100  
 GTAAACATCC GGATGCCACA TACAGCAGGT GCGGCTCTGG TCCCTGGCTT ACCCCTCGAT 2160  
 GCTTGGTAGA CTACCCATAC AGGCTCTGGC ACTACCCCTG TACAATCAAT TACACCATT 2220  
 TCAAGATCAG GATGTTTGTG GCGGGGGTTG AGCACAGGCT TGACGCCGCG TGCAACTGGA 2280  
 CGCGGGGAGA GCGCTGCGAT TTGGACGACA GGGATCGGGC CGAGTTGAGC CCTCTGTTGC 2340  
 TGTCCACTAC GCAATGGCAG GTCCTCCCTT GCTCATTAC AACACTGCCC GCCCTGTCAA 2400  
 CTGGCCTGAT ACATCTCCAC CAGAACATCG TGGACGTGCA GTACCTCTAT GGGTTGAGCT 2460  
 CGGCAGTCAC ATCCTGGGTC ATAAAGTGGG AGTACGTTGT GCTCCTCTTC TTGCTGCTAG 2520  
 CAGATGCTCG CATTGTGCCC TGCTTGTGGA TGATGCTTCT CATATCTCAG GTAGAGGCGG 2580  
 CGCTGGAGAA CTTGATAGTT CTCAACGCTG CTTCCCTAGT CGGGACACAT GGCATCGTCC 2640  
 CCTTCTTCAT CTTTTTTTGT GCAGCTTGGT ACCATAAAGG CAAGTGGGCC CCTGGACTCG 2700  
 CCTATTCCGT CTATGGGATG TGGCCACTGC TCCTGCTTCT CCTGGCGTTG CCCCACGGG 2760  
 CATACGCCTT GGATCAGGAG TTGGCCGCGT CGTGTGGGGC CACGGTCTTC ATCTGCCTAG 2820  
 CGGTGCTCAC TCTATCGCCA TATTACAAAC AGTACATGGC CCGCGGCATC TGGTGGCTGC 2880  
 AGTACATGCT GACCAGAGCA GAGGCGCTCC TACAGGTTTG GGTCCCCCG CTCAACGCC 2940  
 GAGGAGGGCG CGACGGAGTC GTACTGCTCA CGTGTGTGCT CCACCCGCAC TTGCTCTTTG 3000  
 AAATCACCAA GATCATGCTG GCCATTCTCG GGCTTTGTG GATCTTGACG GCCAGTCTGC 3060  
 TCAAGGTACC GTACTTCGTG CGTGCCACG GTCTCATTAG GCTCTGCATG CTGGTGGCA 3120  
 AGACAGCGGG CGGTCAGTAT GTGCAGATGG CTCTGTAAA GCTGGGAGCA TTGCGCGCA 3180

## 【0026】実施例2

## 1c型HCV遺伝子型判定法

(1) 1c型HCVの遺伝子型判定に適したプライマーの選択

1c型HCVの遺伝子型の判定に使用するプライマーを選択するために、実施例1によって明らかになった配列に基づき、1c型HCVに於て最も塩基配列がよく保存されており且つ既存の遺伝子型のHCV株とは相同性が低い遺伝子領域を検索した。その結果、遺伝子判定に適した保存領域は他の遺伝子型と同様にコア領域であることが判明した。これに基づき、コア領域より1c型に特

40 異的な遺伝子断片の増幅に適した配列を有するオリゴ  
 クレオチドプライマーの選択を行った。その結果、第一  
 段階の遺伝子増幅に使用するオリゴヌクレオチドプライ  
 マーとして配列表7および配列表8に記載した#186  
 と#256を選択し、続く第二段階増幅に使用するオリ  
 ゴヌクレオチドプライマーとして配列表9に記載した#  
 104と本発明の#321(配列表6)プライマーを選  
 択した。#186、#256ならびに#104のオリゴ  
 50 ヌクレオチドプライマーは、本発明者らにより既に報告  
 されており(特願平3-307296、特願平4-09  
 3960)公知の遺伝子型判定法で使用されているオリ



13

ゴヌクレオチドプライマーであるが、実施例1によって解明された1c型HCVの配列にもとづき、本発明に於いても利用可能であると判断された。これらのオリゴヌクレオチドプライマーは、本発明に於いて遺伝子型に特異的な保存塩基配列を増幅する為に使用される。他方、本発明の#321オリゴヌクレオチドプライマーは実施例1で解明した1c型についてのみ特異性を有するオリゴヌクレオチドプライマーである。

【0027】(2) 本発明の#321を使用したHCV遺伝子の検出と1c型HCVの遺伝子型判定

各遺伝子型のHCV株(HC-J1:I型、HC-J4:II型、HC-J6:III型、HC-J8:IV型)ならびにHC-G9由来のRNAから、#186のプライマーを使用してcDNAを得た。続いてcDNAを#256ならびに#186のプライマーを使用したPCRを利用して第一段階の増幅を行った。第一段階のPCRは、94℃による変性1分、55℃によるプライマー結合反応1分30秒、72℃によるプライマー伸長反応2分を1サイクルとし35サイクル行った。第二段階の増幅は、型特異オリゴヌクレオチドプライマーである配列番号11~14記載の#296、#133、#134、#135および本発明の#321プライマー、ならびに型共通オリゴヌクレオチドプライマーである配列番号9記載の#104プライマーを用いたPCRにて実施した。第一段階にて増幅された増幅産物の50分の1用量を第二段階増幅のサンプルとしてPCRを行った。反応条件は94℃による変性反応30秒、60℃によるプライマー結合反応30秒、72℃よりなるプライマー伸長反応30秒で、各反応からなる1サイクルを合計30

14

回繰り返した。第二段階終了後、増幅産物を1.5%のNuSieveならびに1.5%のSeaKem(FMC Bioproducts, U. S. A)を用いたアガロース電気泳動し、終了後エチジウムブロマイド染色にてDNAを染色し、各バンドの移動位置より遺伝子型を判定した。各検体の増幅産物の移動度は、各遺伝子型別に設定されたプライマーの位置より予想された移動度、すなわち49、144、174、123、200bp(I型、II型、III型、IV型、1c型)の位置に泳動され、各遺伝子型が判別できることが確認された。また、各検体の電気泳動像には予想される遺伝子型のバンドのみが見られ、別の型に相当するバンドが現われることはなかった。従って、本発明によるHCVの遺伝子型判定法は1c型を含めた全ての遺伝子型について十分に高い特異性を有することが証明された。泳動のパターンを図3に示した。

【0028】

【図3】

【0029】実施例3

各国別の1c型HCVの出現頻度

日本、中華人民共和国、タイ、インドネシア、ニュージーランドのHCV患者検体を対象に1c型の出現頻度を本発明の実施例2の方法を使用して調べた。その結果は表3に示すように1c型HCVはインドネシアに於いてのみ9.9%の頻度で発見され、限局された地域性を有することが判明した。

【0030】

【表3】

15

16

CCTACATTTA CAACCACCTT TCCCCGCTCC AAGACTGGGC TCACAGCGGT CTGCGCGACC 3240  
 TGGCGGTAGC CACTGAACCC GTCATCTTCT CCCGGATGGA AATCAAGACT ATCACCCTGGG 3300  
 GGGCGGATAC TCGGCTTGT GGAGACATCA TCAACGGGCT GCCTGTTTCC GCGCGGAGAG 3360  
 GGAGAGAGGT GTTGCTGGGA CCAGCCGATG CCCTGACTGA CAAAGGATGG AGGCTTTTAG 3420  
 CCCCCATCAC GGCTTACGCC CAGCAGACAC GGGGTCTCTT GGGCTGCATC ATCACCAGCC 3480  
 TCACCGGTCG GGACAAAAAT CAAGTGGAGG GGAAGTCCA GATTGTGTCT ACCGCAACCC 3540  
 AGACGTTCTT GGCTACTTGT GTTAATGGAG TTTGCTGGAC TGTGTATCAT GGGGCCGGAT 3600  
 CGAGGACCAT CGCTTCGGCG TCGGCCCTG TGATCCAGAT GTACACTAAT GTGGACCAGG 3660  
 ATTTGGTGGG CTGGCCAGCG CCTCAGGGAG CGCGCTCCCT GACGCCGTGC ACATGCGGCG 3720  
 CCTCGGATCT GTACTTGGTC ACGAGGCACG CGGACGTCAT TCCAGTGGG CGTGGGGGCG 3780  
 ATAACAGGGG AAGTTTACTA TCTCCCCGGC CAATTTCATA TCTAAAGGGA TCCTCGGGAG 3840  
 GCCCCCTGCT CTGTCCCATG GGACATGCCG TGGGCATTTT CAGGGCCGCG GTGTGCACCC 3900  
 GTGGGGTCGC AAAGGCGGTC GACTTTGTGC CCGTTGAATC CCTAGAGACC ACCATGAGGT 3960  
 CCCCAGTGTT TACCGACAAT TCCAGCCCTC CGACAGTGCC CCAGAGCTAC CAGGTGGCGC 4020  
 ATCTGCACGC TCCACTGGA AGTGGTAAGA GCACGAAGGT GCCGGCCGCC TATGCGGCTC 4080  
 AAGGGTACAA GGTTCCTGTG CTGAACCCGT CTGTTGCTGC CACCCTAGGG TTCGGCGCTT 4140  
 ATATGTCAA GGGCCATGGG ATTGACCCAA ACGTCAGGAC TGGCGTAAGG ACCATTACCA 4200  
 CAGGCTCCCC CATCACCCAC TCCACCTACG GCAAAATTCCT GGCTGACGGT GGGTGTTTAC 4260  
 GAGGTGCGTA TGACATCATA ATATGTGACG AATGTCACTC AGTGGACGCC ACCTCGATTC 4320  
 TAGGCATAGG GACTGTCTTG GACCAAGCGG AGACAGCGGG GGTTAGGCTC ACTATCCTCG 4380  
 CCACCGCTAC ACCACCTGGC TCCGTACCGG TGCCACATTC CAACATCGAG GAAGTTGCAT 4440  
 TGTCCTCTGA GGGGGAGATA CCAITCTATG GTAAGGCCAT CCCCCTAAAT TACATCAAGG 4500  
 GGGGGAGGCA TCTCATTTTC TGTCTTCCA AGAAGAAGTG CGACGAGCTC GCTGCAAAGC 4560  
 TGGTTGGCT GGGCGTCAAC GCAGTGGCTT TTTACCGCGG CCTCGACGTG TCTGTATCC 4620  
 CAACCACAGG AGACGTCGTT GTTGTTGGGA CCGACGCCCT AATGACTGGC TACACCGGCG 4680  
 ATTTGACTC CGTGATAGAC TGCAACACCT GTGTCGTCCA GACAGTCGAT TTCAGCCTAG 4740  
 ACCCTACATT CTCTATTGAG ACTTCCACCG TGCCCCAGGA CGCCGTGTCC CGCTCCCAAC 4800  
 GGAGAGGTAG AACCGGTCGG GGAAGCATG GTATCTACAG ATATGTGTCA CCGGGGAGC 4860  
 GGCCGTCTGG CATGTTTGAC TCCGTGGTCC TCTGTGAGTG CTATGACGCG GGTGTGCTT 4920  
 GGTATGAGCT TACACCGGCC GAGACCACGG TTAGGTTACG GGCATATCTT AACACCCAG 4980  
 GGTGCCCCG GTGCCAGGAC CACTTGGAGT TTTGGGAGAG CGTCTTACC GGCCTCACCC 5040

【0031】

【発明の効果】従来4つの型に分類されていたHCVの遺伝子型は、本発明者らによって新たに5つに分類され、それぞれ特徴的な塩基配列を有することが見出されたことにもとづき、本発明では、型特異的な配列を有する、あるいは型特異的な配列に相補性を有するオリゴヌクレオチドが提供される。当然、これらのオリゴヌクレオチドは遺伝子増幅に於けるプライマーとしてHCV遺伝子の検出と遺伝子レベルの型判定に利用できる。また適当な標識を加えることによりプローブとして単独に、あるいは遺伝子増幅と組み合わせることによって遺伝子レベルの

40 型判定に利用できる。本発明により新たに設定されたオリゴヌクレオチドプライマーを利用することで、従来法では判定不可能であった1c型HCVの遺伝子型も判定可能になることから、より完全な判定技術が確立されたと考えられる。また、本発明によって明らかにされた1c型HCVの遺伝子配列より型特異的な抗原位置の推定が可能であり、本発明の蛋白質またはポリペプチドは遺伝子産物を利用した治療薬開発にも利用できる。

【0032】

【配列表】

配列番号: 1

配列の長さ: 9487

配列の型: 核酸

鎖の数: 一本鎖

トポロジー: 直線状

配列の種類: cDNA to genomic RNA (HC-G9 cDNA)

```

GCCAGCCCC TGATGGGGG GACACTCCGC CATGAATCAC TCCCCTGTGA GGAAGTACTG 60
TCTTCACGCA GAAAGCGTCT AGCCATGGCG TTAGTATGAG TGTCGTGCAG CCTCCAGGAC 120
CCCCCTCCC GGGAGAGCCA TAGTGGTCTG CGGAACCGGT GAGTACACCG GAATTGCCAG 180
GACGACCGGG TCCTTTCTTG GATTAACCCG CTCAATGCCT GGAGATTGG GCGTGCCCC 240
GCAAGACTGC TAGCCGAGTA GTGTTGGGTC GCGAAAGGCC TTGTGGTACT GCCTGATAGG 300
GTGCTTGCGA GTGCCCCGGG AGGTCTCGTA GACCGTGCAC CATGAGCAG AATCCTAAAC 360
CTCAAAGAAA AACCAAACGT AACACCAACC GCGGCCACCA GGACGTTAAG TTCCCGGGTG 420
GCGGCCAGAT CGTTGGCGGA GTTTACTTGT TGCCGCGCAG GGGCCCCAGA GTGGGTGTGC 480
GCGCGACGAG GAAGACTTCC GAGCGGTCCG AACCTCGCGG GAGGCGTCAG CCTATTCCCA 540
AGGCCCGCCG ACCCGAGGGA AGGTCCTGGG CGCAGCCCGG GTACCCTTGG CCCCTCTATG 600
GCAACGAGGG CTGTGGGTGG GCGGGATGGC TCCTGTCCCC CCGCGGCTCT CGGCCTAGTT 660
GGGGCCCTTC TGACCCCGGG CGGAGGTCAC GCAATTTGGG TAAGGTCATC GATACCCTCA 720
CGTGTGGCTT CGCCGACCTC ATGGGGTACA TCCCGCTCGT CGGCGCTCCT CTAGGGGGCG 780
CTGCCAGAGC TCTGGCACAT GGTGTTAGAG TCCTGGAAGA CGGCGTGAAT TACGCAACAG 840
GGAACCTCCC CGGTTGCTCT TTTTCTATCT TCTTGCTCGC TCTTCTATCC TGCCTGACAG 900
TCCCTGCTTC GGCCGTCGGA GTGCGCAACT CTTCGGGGGT GTACCATGTC ACCAATGATT 960
GCCCCAATGC GTCCGTTGTG TACGAGACGG AGAACCTGAT CATGCATCTG CCCGGGTGTG 1020
TGGCCTACGT ACGCGAGGGC AACGCCTCGA GGTGTTGGGT CTCCTTAGT CCCACCGTAG 1080
CCGCCAGGGA TTGCGCGTC CCCGTCACTG AGGTTGGCG TCGTGTGAC TCGATTGTG 1140
GGGCCGCTGC GTTCTGTTG GCTATGTATG TAGGGGACCT ATGCGGCTCC ATCTTCCTG 1200
TTGGCCAGAT CTTACCTTC TCTCCAGGC ACCATTGGAC GACGCAAGAC TGCAATTGCT 1260
CCATCTACCC AGGCCATGTG ACAGGTCATC GAATGGCTTG GGACATGATC ATGAATTGGT 1320

```

19

20

CACCCACTGG CGCCTTAGTG GTGGCAGAGC TACTCCGGAT CCCACAAGCT ATCGTGGATA 1380  
 TGATAGCTGG TGCCCACTGG GGTGTCTAG CGGGCCTGGC ATACTACTCC ATGGTGGGA 1440  
 ACTGGGCTAA GGTGTGGTC GTGCTGCTGC TCTTCGCTGG CATTGACGCA GAGACCCGGG 1500  
 TCACAGGGGG GGCCGCTGGC CACACGGCGT TCGGGTTTGC TAGCTTCCTC GCCCAAGGCG 1560  
 CTAAGCAAAA GATCCAGCTC ATAAATACCA ACGGCAGCTG GCACATCAAC AGAACTGCCT 1620  
 TGAAGTAA TGAAAGCTTG GATACTGGCT GGCTAGCAGG GCTGCTCTAC TACCACAAGT 1680  
 TCAACTCCTC AGGGTGTCCC GAGAGGATGG CTAGTTGCCA ACCTCTTACC GCCTTCGACC 1740  
 AAGGGTGGGG ACCCATCACT CACGAGGGGA ATGCTAGTGA TGACCAGCGG CCATATTGTT 1800  
 GGCATATGC CCTACGCCCG TGTGGCATTG TGCCAGCGAA AAAGGTTTGC GGGCCTGTAT 1860  
 ACTGTTTCAC ACCCAGCCCC GTGGTAGTGG GGACGACGGA CAGAGCCGGC GTTCTACCT 1920  
 ACAGATGGGG TGCCAATGAG ACGGATGTAC TGCTCTCAA CAACTCTAGG CCGCCAATGG 1980  
 GGAATTGGTT TGGGTGTACG TGGATGAATT CTAGTGGCTT CACCAAGACG TCGGGGGCTC 2040  
 CGGCCTGCAA CATCGGCGGG AGCGGGAACA ATACCCTGCT GTGCCAACA GATTGCTTCC 2100  
 GTAAACATCC GGATGCCACA TACAGCAGGT GCGGCTCTGG TCCCTGGCTT ACCCCTCGAT 2160  
 GCTTGGTAGA CTACCCATAC AGGCTCTGGC ACTACCCCTG TACAGTCAAT TACACCATTI 2220  
 TCAAGATCAG GATGTTTGTG GCGGGGGTTG AGCACAGGCT TGACGCCCGG TGCAACTGGA 2280  
 CGCGGGGAGA GCGCTGCGAT TTGGACGACA GGGATCGGGC CGAGTTGAGC CCTCTGTTGC 2340  
 TGTCCACTAC GCAATGGCAG GTCCTCCCTT GTCATTAC AACACTGCCC GCCCTGTCAA 2400  
 CTGGCCTGAT ACATCTCCAC CAGAACATCG TGGACGTGCA GTACCTCTAT GGGTTGAGCT 2460  
 CGGCAGTCAC ATCCTGGGTC ATAAAGTGG AGTACGTTGT GCTCCTCTTC TTGCTGCTAG 2520  
 CAGATGCTCG CATTGTGCC TGCTTGTGGA TGATGCTTCT CATATCTCAG GTAGAGGCGG 2580  
 CGTGAGAGAA CTTGATAGTT CTCAACGCTG CTTCCCTAGT CGGGACACAT GGCAICGICC 2640  
 CCTTCTTCAT CTTTTTTTGT GCAGCTTGGT ACCTAAAAGG CAAGTGGGCC CCTGGACTCG 2700  
 CCTATTCCGT CTATGGGATG TGGCACTGC TCCTGCTTCT CCTGGCGTTG CCCCACGGG 2760  
 CATACGCCCT GGATCAGGAG TTGGCCGCGT CGTGTGGGGC CACGGTCTTC ATCTGCCTAG 2820  
 CGGTGCTCAC TCTATGCCA TATTACAAAC AGTACATGGC CCGCGGCATC TGGTGGCTGC 2880  
 AGTACATGCT GACCAGAGCA GAGGCGCTCC TACAGGTTTG GGTCCCCCG CTCAACGCC 2940  
 GAGGAGGGCG CGACGGAGTC GTACTGCTCA CGTGTGTGCT CCACCCGCAC TTGCTCTTTG 3000  
 AAATACCAA GATCATGCTG GCCATTCTCG GGCCTTTGTG GATCTTGAG GCCAGTCTGC 3060  
 TCAAGGTACC GTACTTCGTG CGTGCCACG GTCTCATTAG GCTCTGCATG CTGGTGGCA 3120  
 AGACAGCGGG CGGTCAATAT GTGCAGATGG CTCTGTAAA GCTGGAGCA TTTGCCGGCA 3180

21

22

CCTACATTA CAACCACCTT TCCCCGCTCC AAGACTGGGC TCACAGCGGT CTGGCGGACC 3240  
 TGGCGGTAGC CACTGAACCC GTCATCTTCT CCCGGATGGA AATCAAGACT ATCACCCTGGG 3300  
 GGGCGGATAC TGGCGCTTGT GGAGACATCA TCAACGGGCT GCCTGTTTCC GCCCGGAGAG 3360  
 GGAGAGAGGT GTTGTGGGA CCAGCCGATG CCCTGACTGA CAAAGGATGG AGGCTTTTAG 3420  
 CCCCCATCAC GGCTTACGCC CAGCAGACAC GGGGTCTCTT GGGCTGCATC ATCACCAGCC 3480  
 TCACCGGTCTG GGACAAAAAT CAAGTGGAGG GGAAGTCCA GATTGTGTCT ACCGCAACCC 3540  
 AGACGTTCTT GGCTACTTGT GTTAATGGAG TTTGCTGGAC TGTGTATCAT GGGGCCGGAT 3600  
 CGAGGACCAT CGCTTCGGCG TCGGGCCCTG TGATCCAGAT GTACACTAAT GTGGACCAGG 3660  
 ATTTGGTGGG CTGGCCAGCG CCTCAGGGAG CGCGCTCCCT GACGCCGTGC ACATGCCGGCG 3720  
 CCTCGGATCT GTACTTGGTC ACGAGGCACG CGGACGTCAT TCCAGTGGCG CGTCGGGGCG 3780  
 ATAACAGGGG AAGTTTACTA TCTCCCCGGC CAATTTTATA TCTAAAGGGA TCCTCGGGAG 3840  
 GCCCCCTGCT CTGTCCCATG GGACATGCCG TGGGCATTTT CAGGGCCGCG GTGTGCACCC 3900  
 GTGGGGTCCG AAAGGCGGTC GACTTTGTGC CCGTTGAATC CCTAGAGACC ACCATGAGGT 3960  
 CCCCAGTGTT TACCGACAAT TCCAGCCCTC CGACAGTGCC CCAGAGCTAC CAGGTGGCGC 4020  
 ATCTGCACGC TCCCACTGGA AGTGGTAAGA GCACGAAGGT GCCGGCCGCC TATGCGGCTC 4080  
 AAGGGTACAA GGTTCCTGTG CTGAACCCGT CTGTTGCTGC CACCCTAGGG TTCGGCGCTT 4140  
 ATATGTCAAA GGCCCATGGG ATTGACCCAA ACGTCAGGAC TGGCGTAAGG ACCATTACCA 4200  
 CAGGCTCCCC CATCACCAC TCCACCTACG GCAAATTCCT GGCTGACGGT GGGTGTTCAG 4260  
 GAGGTGCGTA TGACATCATA ATATGTGACG AATGTCACTC AGTGGACGCC ACCTCGATTG 4320  
 TAGGCATAGG GACTGTCTTG GACCAAGCGG AGACAGCGGG GGTTAGGCTC ACTATCCTCG 4380  
 CCACCGCTAC ACCACCTGGC TCCGTCACCG TGCCACATTC CAACATCGAG GAAGTTGCAT 4440  
 TGTCCACTGA GGGGGAGATA CCATTCTATG GTAAGGCCAT CCCCCTAAAT TACATCAAGG 4500  
 GGGGGAGGCA TCTCATTTTC TGTATTCCA AGAAGAAGTG CGACGAGTC GCTGCAAAGC 4560  
 TGGTTGGCCT GGGCGTCAAC GCAGTGGCCT TTTACCGCGG CCTCGACGTG TCTGTATCC 4620  
 CAACCACAGG AGACGTCGTT GTTGTGGGA CCGACGCCCT AATGACTGGC TACACCGGCG 4680  
 ATTTGCACTC CGTGATAGAC TGCAACACCT GTGTCTCCA GACAGTCGAT TTCAGCCTAG 4740  
 ACCCTACATT CTCTATTGAG ACTTCCACCG TGCCCCAGGA CGCCGTGTCC CGCTCCCAAC 4800  
 GGAGAGGTAG AACCGGTGCG GGAAGCATG GTATCTACAG ATATGTGTCA CCCGGGGAGC 4860  
 GGGCGTCTGG CATGTTTGAC TCCGTGGTCC TCTGTGAGTG CTATGACCGG GGTGTGCTT 4920  
 GGTATGAGCT TACACCGCC GAGACCAGG TTAGGTTACG GGCATATCTT AACACCCAG 4980  
 GGTGCCCCGT GTGCCAGGAC CACTTGGAGT TTTGGGAGAG CGTCTTACC GGCCTCACC 5040

23

24

ACATAGATGC CCACTTCCTG TCTCAGACGA AACAGAGCGG GGAAAATTTC CCCTACCTAG 5100  
 TCGCATACCA AGCCACCGTG TGCCTAGAG CTAAAGCTCC TCCCCGTCA TGGGACCAAA 5160  
 TGTGGAAGTG CTTGATACGG CTCAAGCCCA CCCTCACTGG GGCTACCCCC CTACTATACA 5220  
 GACTGGGTGG TGTGAGAAT GAGATCACCC TAACACACCC AATCACCAAG TACATCATGG 5280  
 CTTGTATGTC GGCTGACCTG GAGGTCGTCA CTAGCACGTG GGTGCTGGTG GGCGGCGTCC 5340  
 TGGCCGCTTT GGCCGCTTAC TGCCTGTCTA CAGGCAGCGT GGTTCATAGTG GGCAGGATAA 5400  
 TCCTAAGCGG GAAGCCGGCA GTCATTCTG ACAGGGAGGT TCTCTACCGA GAGTTTGATG 5460  
 AGATGGAAGA GTGCGCCGCC CACATCCCCT ACCTTGAGCA GGGGATGCAT TTGGCTGAAC 5520  
 AGTTCAAGCA GAAAGCTCTC GGGTTGCTCC AGACAGCATC CAAGCAAGCA GAGACGATCA 5580  
 CTCCCGCTGT CCATACCAAT TGCAAGAAC TCGAGTCCTT CTGGGCTAAG CACATGTGGA 5640  
 ACTTCGTGAG CGGGATACAA TACCTGGCGG GCCTGTCAAC GCTGCCCGGT AATCCCGCTA 5700  
 TAGCGTCGCT GATGTCGTTT ACAGCCGCGG TGACGAGTCC ACTAACCACC CAGCAAACCC 5760  
 TCCTCTTTAA CATCCTGGG GGGTGGGTGG CCGCCAGCT TGCCGCCCA GCTGCCGCCA 5820  
 CTGCTTTCGT CGGCGCTGGT ATTACCGCG CTGTTCATCG CAGTGTGGGC CTAGGGAAGG 5880  
 TCCTAGTGA CATTCTGCT GGCTACGGGG CTGGTGTGGC GGGGGCCCTT GTGGCTTTCA 5940  
 AGATCATGAG CGGGGAGGCC CCCACCGCCG AGGATCTAGT CAACCTTCTG CCTGCCATCC 6000  
 TCTCGCCAGG AGCTCTCGTT GTAGGCGTGG TGTGCGCAGC AATACTACGC CGGCACGTGG 6060  
 GCCCTGGCGA GGGCGCGGTG CAGTGGATGA ACCGACTGAT AGCGTTTGCT TCTCGGGGTA 6120  
 ACCACGTCTC CCCTACACAC TATGTGCCAG AGAGCGACGC GTCAGTCCGT GTCACACATA 6180  
 TCCTCACCAG CCTCACTGTC ACTCAGCTCC TGAAAAGGCT CCACGTGTGG ATAAGCTCAG 6240  
 ATTGCACCGC CCGGTGTGCT GGTTCCTGGC TCAAAGATGT CTGGGACTGG ATATGCGAGG 6300  
 TGCTGAGCGA CTTCAAGAGT TGGCTGAAGG CCAAACCTAT GCCGCAACTG CCCGGGATCC 6360  
 CATTGATATC CTGTCAACGC GGGTACCGTG GGGTCTGGCG GGGCGAAGGC ATCATGCACG 6420  
 CCCGTTGCC GTGTGGAGCC GATATAACTG GTCATGTCAA AAACGGTTCG ATGAGAATCG 6480  
 TCGGCCCTAA GACTTGACGC AACACCTGGC GTGGGTCTGT CCCCATCAAC GCCCACACTA 6540  
 CGGGCCCTTG CACACCCTCC CCAGCGCCGA ACTACACGTT CCGGTTATGG AGGGTGTCTGG 6600  
 CAGAGGAGTA TGTGGAGGTA AGGCGGCTGG GGGATTCCA TTACATCACG GGGGTGACCA 6660  
 CTGATAAGAT CAAGTGTCCA TGCCAGGTCC CCTCGCCCGA GTTCTTCACA GAGGTGGATG 6720  
 GGGTGGCCCT ACATAGGTAC GCGCCCCCT GCAAACCCCT GCTACGGGAT GAGGTGACGT 6780  
 TTAGCATCGG GCTCAATGAA TACTTGGTGG GGTCCAGTT GCCCTGCGAG CCCGAGCCAG 6840  
 ACGTAGCTGT ACTGACATCA ATGCTTACAG ACCCTCCCA CATCACTGCA GAGACGGCGG 6900

25

26

CGCGTAGACT GAATCGGGGG TCTCCCCCT CCCTGGCTAG CTCTTCTGCC AGCCAATTGT 6960  
 CTGCCCGGTC CCTGAAAGCA ACATGTACCA CCCACCATGA CTCTCCAGAC GCTGACCTCA 7020  
 TAACAGCCAA CCTCCTGTGG AGGCAGGAGA TGGGGGGGAA CATTACCAGA GTGGAGTCGG 7080  
 AGAATAAGAT CGTCATCCTG GATTCTTTGG ACCCGCTCGT GGCAGGAGGAG GATGATCGGG 7140  
 AGATTCTGT CCCAGCTGAG ATTCTGCTGA AGTCTAAGAA GTTTCCCCC GCCATGCCCTA 7200  
 TATGGGCACG GCCAGATTAT AATCCTCCCC TTGTGGAACC ATGGAAGCGC CCGGACTACG 7260  
 AACCACCTT AGTTCACGGG TGCCCCCTAC CACCTCCCAA ACCAACTCCG GTGCCGCCAC 7320  
 CCGGAGAAA GAGGACAGTG GTGCTGGATG AGTCTACAGT ATCATCTGCT CTGGCTGAGC 7380  
 TTGCCACTAA GACCTTTGGC AGCTCTACAA CCTCAGGCGT GACAAGTGGT GAAGCGGCCG 7440  
 AATGTCCTCC GCGCGCTTCC TCGACGGTG AACTGGACTC CGAAGCTGAA TCTTACTCCT 7500  
 CCATGCCCC TCTCGAGGGG GAACCGGGG ACCCGATCT CAGCGACGGG TCTTGGTCTA 7560  
 CCGTAAGCAG TGATGGCGGT ACGGAGGACG TCGTGTGCTG CTCGATGTCC TACTCGTGG 7620  
 CCGGCGCCCT AATTACGCCC TGTGCCGCG AGGAAACCAA ACTCCCCATC AACGCACTGA 7680  
 GTAACCTCGT GCTGCGCCAC CACAATTTGG TGTATTCCAC CACCTCTCGC AGCGCTGGCC 7740  
 AGAGGCAGAA AAAAGTCACA TTTGACAGGC TGCAGGTCT GGACGATCAT TACCGGGACG 7800  
 TGCTCAAGGA GGCTAAGGCC AAGGCATCCA CAGTGAAGGC TAAATTGCTA TCCGTAGAGG 7860  
 AGGCATGTAG CCTGACGCCC CCGCACTCCG CCAGATCAA ATTTGGCTAT GGGGCGAAGG 7920  
 ATGTCCGAAG CCATTCCAGT AAGGCCATAC GCCACATTAA CTCCGTGTGG CAAGACCTTC 7980  
 TGGAGGACAA TACAACACCT ATAGACACTA CCATCATGGC AAAGAATGAG GTCTTCTGCG 8040  
 TGAAGCCCGA AAAGGGGGG CGCAAGCCCG CTCGTCTTAT CGTGTACCCC GACCTGGGAG 8100  
 TCGCGGTATG CGAGAAGAGG GCTTTGTATG ACGTAGTCAA ACAGCTCCCC ATTGCCGTGA 8160  
 TGGGAACCTC CTACGGGTTG CAGTACTCAC CAGCGCAGCG GGTGCACTTC CTGCTTAATG 8220  
 CGTGGAAATC AAAGAAAAAT CCTATGGGGT TTTCTTATGA CACCCGTTGC TTTGACTCGA 8280  
 CAGTCACTGA GGCTGATATC CGTACGGAGG AAGACCTCTA CCAATCTTGT GACCTGGTCC 8340  
 CTGAGGCCCG CGCGGCCATA AGGTCTCTCA CAGAGAGGCT TTACATCGGG GGCCCGCTTA 8400  
 CCAACTCTAA GGGACAAAC TGCGGCTATC GCGGATGCCG CCGGAGCGGC GTGCTGACCA 8460  
 CTAGCTGCGG TAACACCATA ACCTGTTATC TCAAGGCCAG TGCAGCCTGT CGAGCTGCAA 8520  
 AGCTCCGGGA CTGCACTATG CTCGTGTGCG GCGACGACCT CGTCGTTATC TGTGAGAGCG 8580  
 CCGGTGTCCA GGAGGACCT GCGAACCTGA GAGCCTTCAC GGAGGCTATG ACCAGGTAAT 8640  
 CCGCCCCCCC GGGAGACCG CCTCAACCAG AATACGACTT GGAGCTTATA ACATCTTGT 8700  
 CCTCCAATGT TTCAGTCGCG CACGACGGCG CTGGCAAAAG GGTCTACTAT CTGACCCGTG 8760

27

28

ATCTGAGAC TCCCCTCGCG CGTGCCGCTT GGGAAACAGC AAGACACACT CCAGTGAAC 8820  
 CCTGGCTAGG CAACATCATC ATGTTTGCCC CCACTCTGTG GGTACGGATG GTCCTTATGA 8880  
 CCCATTTTTT CTCCATACTC ATAGCCCAGG AACACCTTGA AAAGGCTCTA GATTGTGAAA 8940  
 TCTATGGGGC CGTGCACTCC GTCCAACCGT TAGATCTACC TGAAATCATT CAAAGACTCC 9000  
 ACGGCCTCAG CGCGTTCTCG CTCCATAGTT ACTCTCCAGG TGAAATCAAT AGGGTGGCTG 9060  
 CATGCCTCAG GAAACTTGGG GTTCCGCCCT TGCAGGCTTG GAGACACCGG GCCCGGAGCG 9120  
 TCCGGCCAC ACTCCTATCC CAGGGGGGGA GAGCCGCTAT ATGCGGTAAG TATCTCTTCA 9180  
 ACTGGGCGGT GAAAACCAAA CTCAAACCTCA CTCCATTACC GTCCGCGTCT CAGTTGGACT 9240  
 TGTCCAATTG GTTCACGGGC GGTACAGCG GGGGAGACAT TTATCACAGC GTGTCTCATG 9300  
 TCCGGCCCCG CTGGTTCCTC TGGTGCCTAC TCCTACTTTC AGTGGGGGTA GGCATTTACC 9360  
 TCCTTCCCAA CCGGTAGACG GTTGGGCAAC CACTCCAGGC CTTTAGGCCC TGTTTAAACA 9420  
 CTCCAGGCCT TTAGGCCCCG TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT 9480  
 TTTTTT 9487

[0033]



配列番号: 2

配列の長さ: 1765

配列の型: 核酸

鎖の数: 一本鎖

トポロジー: 直線状

配列の種類: cDNA to genomic RNA (YS117-5' cDNA)

```

CCATGGCGTT AGTATGAGTG TCGTGCAGCC TCCAGGACCC CCCCTCCCGG GAGAGCCATA 60
GTGGTCTGCG GAACCGGTGA GTACACCGGA ATTGCCAGGA CGACCGGGTC CTTTCTTGGA 120
TCAACCCGCT CAATGCCTGG AGATTGGGGC GTGCCCCCGC AAGACTGCTA GCCGAGTAGT 180
GTTGGGTGCG GAAAGGCCTT GTGGTACTGC CTGATAGGGT GCTTGCGAGT GCCCCGGGAG 240
GTCTCGTAGA CCGTGCACCA TGAGCACAAA TCCTAACCT CAAAGAAAAA CCAAACGTAA 300
CACCAACCGC CGCCACAGG ACGTTAAGT CCCGGGTGGC GGCCAGATCG TTGGCGGAGT 360
TTACTTGTG CCGCGCAGGG GCCCCAGAGT GGGTGTGCG GCGACGAGGA AGACTTCCGA 420
GCGGTGCGAA CCTCGCGGAA GGGTGCAGCC TATCCCAAG GCCCGCCGAC CCGAGGGTAG 480
GTCCTGGGCG CAGCCCGGGT ACCCTTGGCC CCTCTATGG AACGAGGGCT GCGGGTGGGC 540
GGGATGGCTC CTGTCCCCC GCGGCTCTCG GCCTAGTTGG GGCCCCACTG ACCCCCGGCG 600
GAGGTCACGC AATTGGGTA AGGTCATCGA TACTCTCAG TGTGGCTTCG CCGACCTCAT 660
GGGTACATC CCGCTCGTC GTGCTCCTCT AGGGGGCGCT GCCAGAGCTC TGGCACACGG 720
TGTTAGAGTT CTGGAAGAC GCGTGAAC TAACACAGG AACCTTCCTG GTTGCTCCTT 780
TTCTATCTT TTGCTCGCTC TTCTATCCTG CCTGACAGTC CCTGCTTCGG CCGTCGAAGT 840
GCGCAACTCA TCAGGGGTGT ACCATGTAC CAATGATTGC CCCAATGCGT CCGTTGTGTA 900
CGAGACAGAG AGCCTGATCA TGCATCTGCC CGGGTGTGTG CCCTGCGTAC GCGAGGGCAA 960
CGCCTCGAGG TGCTGGGTCT CCCTTAGCCC TACCATTGCC GCTAAGGATC CGAGCGTCCC 1020
CGTCAGTGAG ATTCGGCGTC ACGTCGACTT GATCGTCGGG GCCGCCGCGT TCTGTTGCGC 1080
TATGTATGTA GGGGACCTAT GCGGCTCCAT CTTCCTCGTT GGCCAGATTI TCACCTCTC 1140
TCCAGGGCGC CATTGGACGA CGCAGGACTG TAATTGCTCC ATCTACCCGG GCCATGTGAC 1200
AGGTCATCGA ATGGCTTGGG ACATGATGAT GAATTGGTCA CCCACTGGCG CCCTAGTGAT 1260

GGCGCAGCTA CTCCGGATCC CACAAGCTGT CGTGGATATG ATAGCCGGTG CCCACTGGGG 1320
TGTCTAGCG GGCCTGGCAT ACTACTCCAT GGTGGGGAAC TGGGCTAAGG TTGTGGTTGT 1380
GCTGTTGCTC TTCGCTGGCG TCGACGCGGA CACCCAGGTC ACAGGAGGCA GCGCTGCCTA 1440
TGATGCGCGC GGACTTGCTT CCTTTTTCAC CCCAGGCCCT AAGCAAAACA ICCAGCTCAT 1500
AAATACCAAC GGCAGCTGGC ACATCAACAG GACCGCCTTG AACTGTAATG AAAGCCTGAA 1560
CACCGGCTGG GTAGCAGGCC TGTCTACTA TCACAAATC AACTCCTCGG GGTGTCCTGA 1620
GAGGATGGCT AGTTGCCAGC CCTCACCAG CTTTGACCAA GGGTGGGGAC CCATCACTTA 1680
CGAGGGGAAT GCTAGCGGCG ACCAACGGCC ATATTGCTGG CACTATGCCC CACGCCCCGTG 1740
CGGTATTGTG CCGCGGAGAG AGGTT 1765

```

[0034]

配列番号: 3

配列の長さ: 1191

配列の型: 核酸

鎖の数: 一本鎖

トポロジー: 直線状

配列の種類: cDNA to genomic RNA (YS117-3' cDNA)

```

AACAGTCACT GAGGCTGATA TCCGTACGGA GGAAGACCTC TACCAATCTT GTGACCTGGT   60
CCCCGAGGCC CGCACGGCCA TAAGGTCTCT CACAGAGAGG CTTTACATCG GGGGCCCCCT   120
TACCAATTCC AAGGGACAAA ACTGCGGCTA TCGGCGATGC CGTGCAAGCG GCGTGCTGAC   180
CACTAGCTGC GGTAAACACCA TAACCTGTTA TCTCAAGGCC AGCGCAGCCT GTCGAGCTGC   240
AAAGCTCCAG GACTGCACCA TGCTCGTGTG CCGCGACGAC CTCGTCGTTA TCTGTGAGAG   300
CGCCGGTGTG CAGGAGGACG CTGCGAGCCT GAGAGCCTTC ACGGAGGCTA TGACCAGGTA   360
CTCCGCCCCC CCGGGAGACC CGCCTCAACC AGAATACGAC TTGGAGCTCA TAACATCCTG   420
TTCCTCCAAC GTGTCAGTCG CGCAGGACGG CTCTGGCAAA AGGGTCTACT ATCTGACCCG   480
TGATCCTGAG ACTCCCTCG CGCGTGCCGC TTGGGAGACA GCAAGACACA CTCCAGTGAA   540
CTCCTGGCTA GGCAACATCA TCATGTTTGC CCCCACTCTG TGGGTACGGA TGGTTCTTAT   600
GACCCATTTT TTTTCCATAC TCATAGCCCA GGAGCACCTT GAAAAGGCTC TAGATTGTGA   660
AATCTATGGA GCCGTACACT CCGTCCAACC GCTGGACCTA CCTGAAATCA TTCAAAGACT   720
CCACGGCCTC AGCGCGTTTT CGCTCCACAG TTA CTCTCCA GGTGAAATCA ATAGGGTGGC   780
TGCATGCCTC AGAAAACCTG GGGTTCCGCC CTGCGAGCT TGGAGACACC GGGCCCGGAG   840
CGTCCGCGCC AACTCCTAT CCCAGGGGGG GAAAGCTGCT ATATGCGGTA AGTACCTCTT   900
CAACTGGGCG GTGAAAACCA AACTCAAAC TACTCCATTA CCGTCCGCGT CTCAGTTGGA   960
CTTGTTCAAT TGGTTCACGG GCGGCTACAG CCGGGGAGAC ATTTATCACA GCGTGTCTCA 1020
TGTCCGCCCC CGTTGGTTCT TCTGGTGCCT ACTCCTACTT TCAGTGGGGG TAGGCATCTA 1080
TCTCCTCCC AACCGATAGA CCGTTGGGCA ATCACTCCTA GCCTTTAGGC CTTATTTAAA 1140
CACTCCAGGC CTTTAGGCCG TGTITTTTTT TTTTTTTTTT TTTTTTTTTT T 1191

```

[0035]

配列番号: 4

配列の長さ: 1765

配列の型: 核酸

鎖の数: 一本鎖

トポロジー: 直線状

配列の種類: cDNA to genomic RNA (SR037-5' cDNA)

```

CCATGGCGTT AGTATGAGTG TCGTGCAGCC TCCAGGACCC CCCCTCCCGG GAGAGCCATA 60
GTGGTCTGCG GAACCGGTGA GTACACCGGA ATTGCCAGGA CGACCGGGTC CTTTCTTGGA 120
TTAACCCGCT CAATGCCTGG AGATTGGGC GTGCCCCCGC AAGACTGCTA GCCGAGTAGT 180
GTTGGGTGCG GAAAGGCCCT GTGGTACTGC CTGATAGGGT GCTTGCGAGT GCCCCGGGAG 240
GTCTCGTAGA CCGTGCACCA TGAGCAGGAA TCCTAAACCT CAAAGAAAAA CCAAACGTAA 300
CACCAACCGC CGCCACAGG ACGTCAAGTT CCCGGGTGGC GGCCAGATCG TTGGCGGAGT 360
TTACTTGTG CCGCGCAGGG GCCCCAGAAT GGGTGTGCG GCGACGAGGA AGACTTCCGA 420
GCGGTGCGAA CCTCGCGGAA GCGGTGAGCC TATTCCCAAG GCCCGCCGAC CCGAGGGTAG 480
GTCCTGGGCG CAGCCCGGGT ACCCTTGGCC CCTCTATGGT AACGAGGGCT GTGGGTGGGC 540
GGGATGGCTT CTGTCCCCC GCGGTTCGGC GCCTAGTTGG GGCCCTCTG ACCCCGGCG 600
GAGGTACGCG AACTTGGTA AGGTCATCGA TACCCTCAG TGTGGCTTCG CCGACCTCAT 660
GGGGTACATC CCGCTCGTCG GTGCTCCTTT AGGGGGCGCT GCCAGAGCTC TGGCGCATGG 720
TGTGAGAGTT CTGGAAGACG GCGTGAATTA TGCAACAGGG AACCTTCCCG GTTGCTCTTT 780
TTCTATCTTC TTGCTTGCCC TTCTATCCTG CCTGACAGTC CCTGCTTCGG CCGTCGGAGT 840
GCGCAACTCT TCGGGGGTGT ACCATGTAC CAATGATTGC CCCAATGGT CTGTTGTGTA 900
CGAGACAGAG AGCCTGATCA TACATCTGCC CGGGTGTGTG CCCTGCGTAC GCGAGGGCAA 960
CGCCTCGAGG TGCTGGGTCT CCCTAGTCC TACTGTTGCC GCTAAGGATC CGAGCGTCCC 1020
CGTCAGTGAG ATTCGACGCC ATGTGACCT GATTGTGGG GCCGCTGCGT TCTGTTGGC 1080
TATGTACGTA GGGGACCTAT GCGGCTCCAT CTTCCTCGTT GGCCAGATTT TCACCCTCTC 1140
TCCAGGCGT CACTGGACGA CGCAGGACTG TAATTGTTCC ATCTACCCAG GCCATGTGAC 1200
AGGTCATCGA ATGGCTTGGG ACATGATGAT GAATGGTCA CCTACTGGCG CCCTAGTGGT 1260
GGCGCAGCTA CTCCGGATCC CACAAGCTGT CGTGGATATG ATAGCCGGTG CCCACTGGGG 1320
TGTCCTAGCG GGCCTGGCAT ACTATTCAT GGTGGGGAAC TGGGCTAAGG TTGTGGTTGT 1380
GCTGCTACTT TTTGCTGGCG TCGATGCAGA GACCCAGGTC TCAGGAGGCT CCGCTGCCCA 1440
AACCACGTAC GGTCTTACTG CCCTCTTCAG GACAGGCCCT AATCAAAAAA TCCAGCTCAT 1500
AAATACCAAC GGCAGCTGGC ATATCAACAG GACCGCCTTG AACTGTAATG AGAGCTTGCA 1560
CACCGGCTGG CTGGCAGCGC TGTCTACAC CCACAAGTC AACTCTTCGG GGTGTTTGA 1620
GAGGATGGCC AGTTGCCAGC CTCTTTCGGC CTTCGACCAA GGGTGGGGAC CCATCACTTA 1680
CGGGGGGAAT GCTAGCGACG GCCAACGGCC ATATTGCTGG CACTATGCCC CACGCCCCGTG 1740
CGGTATTGTG CCGGCGAGAG AGGTT 1765

```

【0036】

配列番号：5

配列の長さ：1179

配列の型：核酸

鎖の数：一本鎖

トポロジー：直線状

配列の種類：cDNA to genomic RNA (SR037-3' cDNA)

AACAGTCACT GAGGCTGATA TCCGCACGGA GGAAGACCTC TACCAATCTT GTGACCTGGT 60  
 CCCTGAGGCC CGCACGGCCA TAAGGTCCCT CACAGAGAGG CTTTACATCG GGGGCCCCGT 120  
 TACCAATTCT AAGGGACAAA ACTGCGGCTA TCGGCGATGC CGCGCAAGCG GCGTGCTGAC 180  
 CACTAGCTGC GGTAACACCA TAACCTGTTA TCTCAAGGCC AGTGCAGCCT GTCGAGCTGC 240  
 AAAGCTCCGG GACTGCACTA TGCTCGTGTG CGGCGATGAC CTTGTCTGTA TCTGTGAGAG 300  
 CGCCGGTGTG CAGGAGGACG CTGCGAGCCT GAGAGCCTTC ACGGAGGCTA TGACCAGGTA 360  
 CTCTGCCCCC CCGGGAGACC CGCCTCAACC AGAATACGAC TTGGAGCTTA TAACATCCTG 420  
 TTCCTCAAT GTGTCACTG CGCACGACGG CGCTGGCAAA AGGGTCTACT ATCTGACCCG 480  
 TGATCCTGAG ACCCCCTCG CGCGTGCCGC TTGGGAGACA GCAAGACACA CTCCAGTGAA 540  
 CTCCTGGCTA GGCAACATCA TTATGTTTGC CCCCACTTTG TGGGTACGGA TGGTCTCAT 600  
 GACCCATTTT TTCTCCATAC TCATAGCCCA GGAGCACCTT GAAAAGGCTC TAGATTGTGA 660  
 AATCTATGGA GCCGTACACT CCATCCAACC GCTGGACCTA CCTGAAATCA TTCAAAGACT 720  
 CCACGGCCTC AGCGCGTTTT CGCTCCACAG TTA CTCTCCA GGTGAAATCA ATAGGGTGGC 780  
 TGCATGCCTC AGAAAACTTG GGGTTCCGCC CTTGCGAGCT TGGAGACACC GGGCCCCGAG 840  
 CGTCCGCGCC ACATCTCTAT CCCAGGGGGG GAAAGCCGCT ATATGCGGTA AGTACCTCTT 900  
 CAACTGGGCG GTGAAAACCA AACTCAAACCT CACTCCATTA CCGTCCGCGT CTCAGTTGGA 960  
 CTTGTCCAAT TGGTTCACGG GCGGCTACAA CGGGGGAGAC ATTTATCACA GCGTGTCTCG 1020  
 TGTCCGGCCC CGTTGGTTCT TCTGGTGCTT ACTCCTACTC TCAGTGGGGG TAGGCATCTA 1080  
 TCTCCTTCCC AACCGATAGA CGGTTGGGTA ATCACTCCAA GCCTTTAGGC CCTTTTTAAA 1140  
 CACTCCAGGC CTTTTGGCCC TGTTTTTTTT TTTTTTTTT 1179

【0037】配列番号：6

配列の長さ：20

配列の型：核酸

鎖の数：一本鎖

トポロジー：直線状

配列の種類：cDNA to genomic RNA

特徴を決定した方法：E

配列（#321）

AACCTCCGCC GGGGATCAGA 20

【0038】配列番号：7

配列の長さ：20

配列の型：核酸

鎖の数：一本鎖

40 トポロジー：直線状

配列の種類：cDNA to genomic RNA

特徴を決定した方法：E

配列（#186）

AYGTACCCCA YGAGRTCCGC 20

（YはTまたはC。RはGまたはA）

【0039】配列番号：8

配列の長さ：20

配列の型：核酸

鎖の数：一本鎖

50 トポロジー：直線状

37

配列の種類: cDNA to genomic RNA  
 特徴を決定した方法: E  
 配列 (#256)  
 CGCGCGMCNA GGAARRCTTC 20  
 (MはAまたはC。NはA, T, CまたはG。RはAまたはG)  
 【0040】配列番号: 9  
 配列の長さ: 20  
 配列の型: 核酸  
 鎖の数: 一本鎖  
 トポロジー: 直線状  
 配列の種類: cDNA to genomic RNA  
 特徴を決定した方法: E  
 配列 (#104)  
 AGRAARRCTT CSGAGCGRTC 20  
 (RはGまたはA。SはCまたはG。)  
 【0041】配列番号: 10  
 配列の長さ: 20  
 配列の型: 核酸  
 鎖の数: 一本鎖  
 トポロジー: 直線状  
 配列の種類: cDNA to genomic RNA  
 特徴を決定した方法: E  
 配列 (#132)  
 YRCCTTGGGC ATAGGCTGAC 20  
 (YはTまたはC。RはGまたはA。)  
 【0042】配列番号: 11  
 配列の長さ: 20  
 配列の型: 核酸  
 鎖の数: 一本鎖  
 トポロジー: 直線状  
 配列の種類: cDNA to genomic RNA  
 特徴を決定した方法: E  
 配列 (#133)

38

GARCCAWCCT GCCCAYCCYA 20  
 (RはGまたはA。WはTまたはA。YはCまたはT。)  
 【0043】配列番号: 12  
 配列の長さ: 20  
 配列の型: 核酸  
 鎖の数: 一本鎖  
 トポロジー: 直線状  
 配列の種類: cDNA to genomic RNA  
 10 特徴を決定した方法: E  
 配列 (#134)  
 CCAARAGGGA CGGGARCCTC 20  
 (RはGまたはA。)  
 【0044】配列番号: 13  
 配列の長さ: 20  
 配列の型: 核酸  
 鎖の数: 一本鎖  
 トポロジー: 直線状  
 配列の種類: cDNA to genomic RNA  
 20 特徴を決定した方法: E  
 配列 (#135)  
 RCCYTCGTTT CCRTACAGRG 20  
 (RはGまたはA。YはCまたはT。)  
 【0045】配列番号: 14  
 配列の長さ: 20  
 配列の型: 核酸  
 鎖の数: 一本鎖  
 トポロジー: 直線状  
 配列の種類: cDNA to genomic RNA  
 30 特徴を決定した方法: E  
 配列 (#296)  
 GGATAGGCTG ACGTCTACCT 20  
 【0046】

配列番号: 15

配列の長さ: 3011

配列の型: アミノ酸

配列の種類: 蛋白質 (HC-G9 amino acid)

Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr		
	5	10
Asn Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gly Gln Ile		
	20	25
Val Gly Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Val Gly		
	35	40
Val Arg Ala Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro Arg Gly		
	50	55
Arg Arg Gln Pro Ile Pro Lys Ala Arg Arg Pro Glu Gly Arg Ser		
	65	70
Trp Ala Gln Pro Gly Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly		
	80	85
Cys Gly Trp Ala Gly Trp Leu Leu Ser Pro Arg Gly Ser Arg Pro		
	95	100
Ser Trp Gly Pro Ser Asp Pro Arg Arg Arg Ser Arg Asn Leu Gly		
	110	115
Lys Val Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu Met Gly		
	125	130
Tyr Ile Pro Leu Val Gly Ala Pro Leu Gly Gly Ala Ala Arg Ala		
	140	145
Leu Ala His Gly Val Arg Val Leu Glu Asp Gly Val Asn Tyr Ala		
	155	160
Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile Phe Leu Leu Ala		
	170	175
Leu Leu Ser Cys Leu Thr Val Pro Ala Ser Ala Val Gly Val Arg		

41	185	190	42	195
Asn Ser Ser Gly Val Tyr His Val Thr		Asn Asp Cys Pro Asn Ala		
200		205		210
Ser Val Val Tyr Glu Thr Glu Asn Leu Ile Met His Leu Pro Gly				
215		220		225
Cys Val Pro Tyr Val Arg Glu Gly Asn Ala Ser Arg Cys Trp Val				
230		235		240
Ser Leu Ser Pro Thr Val Ala Ala Arg Asp Ser Arg Val Pro Val				
235		240		245
Ser Glu Val Arg Arg Arg Val Asp Ser Ile Val Gly Ala Ala Ala				
250		255		260
Phe Cys Ser Ala Met Tyr Val Gly Asp Leu Cys Gly Ser Ile Phe				
265		270		275
Leu Val Gly Gln Ile Phe Thr Phe Ser Pro Arg His His Trp Thr				
280		285		290
Thr Gln Asp Cys Asn Cys Ser Ile Tyr Pro Gly His Val Thr Gly				
295		300		305
His Arg Met Ala Trp Asp Met Met Met Asn Trp Ser Pro Thr Gly				
315		320		325
Ala Leu Val Val Ala Gln Leu Leu Arg Ile Pro Gln Ala Ile Val				
330		335		340
Asp Met Ile Ala Gly Ala His Trp Gly Val Leu Ala Gly Leu Ala				
345		350		355
Tyr Tyr Ser Met Val Gly Asn Trp Ala Lys Val Val Val Val Leu				
360		365		370
Leu Leu Phe Ala Gly Val Asp Ala Glu Thr Arg Val Thr Gly Gly				
375		380		385
Ala Ala Gly His Thr Ala Phe Gly Phe Ala Ser Phe Leu Ala Pro				
390		395		400
Gly Ala Lys Gln Lys Ile Gln Leu Ile Asn Thr Asn Gly Ser Trp				
405		410		415

43	44
His Ile Asn Arg Thr Ala Leu Asn Cys Asn Glu Ser Leu Asp Thr	
420	425 430
Gly Trp Leu Ala Gly Leu Leu Tyr Tyr His Lys Phe Asn Ser Ser	
435	440 445
Gly Cys Pro Glu Arg Met Ala Ser Cys Gln Pro Leu Thr Ala Phe	
450	455 460
Asp Gln Gly Trp Gly Pro Ile Thr His Glu Gly Asn Ala Ser Asp	
465	470 475
Asp Gln Arg Pro Tyr Cys Trp His Tyr Ala Leu Arg Pro Cys Gly	
480	485 490
Ile Val Pro Ala Lys Lys Val Cys Gly Pro Val Tyr Cys Phe Thr	
495	500 505
Pro Ser Pro Val Val Val Gly Thr Thr Asp Arg Ala Gly Val Pro	
510	515 520
Thr Tyr Arg Trp Gly Ala Asn Glu Thr Asp Val Leu Leu Leu Asn	
525	530 535
Asn Ser Arg Pro Pro Met Gly Asn Trp Phe Gly Cys Thr Trp Met	
540	545 550
Asn Ser Ser Gly Phe Thr Lys Thr Cys Gly Ala Pro Ala Cys Asn	
555	560 575
Ile Gly Gly Ser Gly Asn Asn Thr Leu Leu Cys Pro Thr Asp Cys	
570	575 580
Phe Arg Lys His Pro Asp Ala Thr Tyr Ser Arg Cys Gly Ser Gly	
585	590 595
Pro Trp Leu Thr Pro Arg Cys Leu Val Asp Tyr Pro Tyr Arg Leu	
600	605 610
Trp His Tyr Pro Cys Thr Val Asn Tyr Thr Ile Phe Lys Ile Arg	
615	620 625
Met Phe Val Gly Gly Val Glu His Arg Leu Asp Ala Ala Cys Asn	
630	635 640
Trp Thr Arg Gly Glu Arg Cys Asp Leu Asp Asp Arg Asp Arg Ala	



45	650	655	660
Glu Leu Ser Pro Leu Leu Leu Ser Thr Thr Gln Trp Gln Val Leu			
665	670	675	
Pro Cys Ser Phe Thr Thr Leu Pro Ala Leu Ser Thr Gly Leu Ile			
680	685	690	
His Leu His Gln Asn Ile Val Asp Val Gln Tyr Leu Tyr Gly Leu			
695	700	705	
Ser Ser Ala Val Thr Ser Trp Val Ile Lys Trp Glu Tyr Val Val			
710	715	720	
Leu Leu Phe Leu Leu Leu Ala Asp Ala Arg Ile Cys Ala Cys Leu			
725	730	735	
Trp Met Met Leu Leu Ile Ser Gln Val Glu Ala Ala Leu Glu Asn			
740	745	750	
Leu Ile Val Leu Asn Ala Ala Ser Leu Val Gly Thr His Gly Ile			
755	760	765	
Val Pro Phe Phe Ile Phe Phe Cys Ala Ala Trp Tyr Leu Lys Gly			
770	775	780	
Lys Trp Ala Pro Gly Leu Ala Tyr Ser Val Tyr Gly Met Trp Pro			
785	790	795	
Leu Leu Leu Leu Leu Leu Ala Leu Pro Gln Arg Ala Tyr Ala Leu			
800	805	810	
Asp Gln Glu Leu Ala Ala Ser Cys Gly Ala Thr Val Phe Ile Cys			
810	815	820	
Leu Ala Val Leu Thr Leu Ser Pro Tyr Tyr Lys Gln Tyr Met Ala			
830	835	840	
Arg Gly Ile Trp Trp Leu Gln Tyr Met Leu Thr Arg Ala Glu Ala			
845	850	855	
Leu Leu Gln Val Trp Val ProP ro Leu Asn Ala Arg Gly Gly Arg			
860	865	870	
Asp Gly Val Val Leu Leu Thr Cys Val Leu His Pro His Leu Leu			
885	890	895	

47	48
Phe Glu Ile Thr Lys Ile Met Leu Ala Ile Leu Gly Pro Leu Trp	
900	905 910
Ile Leu Gln Ala Ser Leu Leu Lys Val Pro Tyr Phe Val Arg Ala	
915	920 925
His Gly Leu Ile Arg Leu Cys Met Leu Val Arg Lys Thr Ala Gly	
930	935 940
Gly Gln Tyr Val Gln Met Ala Leu Leu Lys Leu Gly Ala Phe Ala	
945	950 955
Gly Thr Tyr Ile Tyr Asn His Leu Ser Pro Leu Gln Asp Trp Ala	
960	965 970
His Ser Gly Leu Arg Asp Leu Ala Val Ala Thr Glu Pro Val Ile	
975	980 985
Phe Ser Arg Met Glu Ile Lys Thr Ile Thr Trp Gly Ala Asp Thr	
990	995 1000
Ala Ala Cys Gly Asp Ile Ile Asn Gly Leu Pro Val Ser Ala Arg	
1005	1010 1015
Arg Gly Arg Glu Val Leu Leu Gly Pro Ala Asp Ala Leu Thr Asp	
1020	1025 1030
Lys Gly Trp Arg Leu Leu Ala Pro Ile Thr Ala Tyr Ala Gln Gln	
1035	1040 1045
Thr Arg Gly Leu Leu Gly Cys Ile Ile Thr Ser Leu Thr Gly Arg	
1050	1055 1060
Asp Lys Asn Gln Val Glu Gly Glu Val Gln Ile Val Ser Thr Ala	
1065	1070 1075
Thr Gln Thr Phe Leu Ala Thr Cys Val Asn Gly Val Cys Trp Thr	
1080	1085 1090
Val Tyr His Gly Ala Gly Ser Arg Thr Ile Ala Ser Ala Ser Gly	
1095	1100 1105
Pro Val Ile Gln Met Tyr Thr Asn Val Asp Gln Asp Leu Val Gly	
1110	1115 1120
Trp Pro Ala Pro Gln Gly Ala Arg Ser Leu Thr Pro Cys Thr Cys	

49	50
1025	1030
Gly Ala Ser Asp Leu Tyr Leu Val Thr Arg His Ala Asp Val Ile	1035
1040	1045
Pro Val Arg Arg Arg Gly Asp Asn Arg Gly Ser Leu Leu Ser Pro	1050
1055	1060
Arg Pro Ile Ser Tyr Leu Lys Gly Ser Ser Gly Gly Pro Leu Leu	1065
1070	1075
Cys Pro Met Gly His Ala Val Gly Ile Phe Arg Ala Ala Val Cys	1080
1085	1090
Thr Arg Gly Val Ala Lys Ala Val Asp Phe Val Pro Val Glu Ser	1095
1100	1105
Ileu Glu Thr Thr Met Arg Ser Pro Val Phe Thr Asp Asn Ser Ser	1110
1115	1120
Pro Pro Thr Val Pro Gln Ser Tyr Gln Val Ala His Leu His Ala	1125
1130	1135
Pro Thr Gly Ser Gly Lys Ser Thr Lys Val Pro Ala Ala Tyr Ala	1140
1145	1150
Ala Gln Gly Tyr Lys Val Leu Val Leu Asn Pro Ser Val Ala Ala	1155
1160	1165
Thr Leu Gly Phe Gly Ala Tyr Met Ser Lys Ala His Gly Ile Asp	1170
1175	1180
Pro Asn Val Arg Thr Gly Val Arg Thr Ile Thr Thr Gly Ser Pro	1185
1190	1195
Ile Thr His Ser Thr Tyr Gly Lys Phe Leu Ala Asp Gly Gly Cys	1200
1205	1210
Ser Gly Gly Ala Tyr Asp Ile Ile Ile Cys Asp Glu Cys His Ser	1215
1225	1230
Val Asp Ala Thr Ser Ile Leu Gly Ile Gly Thr Val Leu Asp Gln	1235
1240	1245
Ala Glu Thr Ala Gly Val Arg Leu Thr Ile Leu Ala Thr Ala Thr	1250
1255	1260
	1265

51	52
Pro Pro Gly Ser Val Thr Val Pro His Ser Asn Ile Glu Glu Val	
1270	1275 1280
Ala Leu Ser Thr Glu Gly Glu Ile Pro Phe Tyr Gly Lys Ala Ile	
1385	1290 1295
Pro Leu Asn Tyr Ile Lys Gly Gly Arg His Leu Ile Phe Cys His	
1300	1305 1310
Ser Lys Lys Lys Cys Asp Glu Leu Ala Ala Lys Leu Val Gly Leu	
1315	1320 1325
Gly Val Asn Ala Val Ala Phe Tyr Arg Gly Leu Asp Val Ser Val	
1330	1335 1340
Ile Pro Thr Thr Gly Asp Val Val Val Val Ala Thr Asp Ala Leu	
1345	1350 1355
et Thr Gly Tyr Thr Gly Asp Phe Asp Ser Val Ile Asp Cys Asn	
1360	1365 1370
Thr Cys Val Val Gln Thr Val Asp Phe Ser Leu Asp Pro Thr Phe	
1375	1380 1385
Ser Ile Glu Thr Ser Thr Val Pro Gln Asp Ala Val Ser Arg Ser	
1390	1395 1400
Gln Arg Arg Gly Arg Thr Gly Arg Gly Lys His Gly Ile Tyr Arg	
1405	1410 1415
Tyr Val Ser Pro Gly Glu Arg Pro Ser Gly Met Phe Asp Ser Val	
1420	1525 1430
Val Leu Cys Glu CysTyr Asp Ala Gly Cys Ala Trp Tyr Glu Leu	
1435	1440 1445
Thr Pro Ala Glu Thr Thr Val Arg Leu Arg Ala Tyr Leu Asn Thr	
1450	1450 1455
Pro Gly Leu Pro Val Cys Gln Asp His Leu Glu Phe Trp Glu Ser	
1465	1470 1475
Val Phe Thr Gly Leu Thr His Ile Asp Ala His Phe Leu Ser Gln	
1480	1485 1490
Thr Lys Gln Ser Gly Glu Asn Phe Pro Tyr Leu Val Ala Tyr Gln	

53	1495	1500	54	1505
	Ala Thr Val Cys Ala Arg Ala Lys Ala Pro Pro Pro Ser Trp Asp			
	1510	1515		1520
	Gln Met Trp Lys Cys Leu Ile Arg Leu Lys Pro Thr Leu Thr Gly			
	1525	1530		1535
	Ala Thr Pro Leu Leu Tyr Arg Leu Gly Gly Val Gln Asn Glu Ile			
	1540	1545		1550
	Thr Leu Thr His Pro Ile Thr Lys Tyr Ile Met Ala Cys Met Ser			
	1555	1560		1565
	Ala Asp Leu Glu Val Val Thr Ser Thr Trp Val Leu Val Gly Gly			
	1570	1575		1580
	Val Leu Ala Ala Leu Ala Ala Tyr Cys Leu Ser Thr Gly Ser Val			
	1585	1590		1595
	Val Ile Val Gly Arg Ile Ile Leu Ser Gly Lys Pro Ala Val Ile			
	1600	1605		1610
	Pro Asp Arg Glu Val Leu Tyr Arg Glu Phe Asp Glu Met Glu Glu			
	1615	1620		1625
	Cys Ala Ala His Ile Pro Tyr Leu Glu Gln Gly Met His Leu Ala			
	1730	1635		1640
	Glu Gln Phe Lys Gln Lys Ala Leu Gly Leu Leu Gln Thr Ala Ser			
	1645	1650		1655
	Lys Gln Ala Glu Thr Ile Thr Pro Ala Val His Thr Asn Trp Gln			
	1660	1665		1670
	Lys Leu Glu Ser Phe Trp Ala Lys His Met Trp Asn Phe Val Ser			
	1675	1680		1685
	Gly Ile Gln Tyr Leu Ala Gly Leu Ser Thr Leu Pro Gly Asn Pro			
	1690	1695		1700
	Ala Ile Ala Ser Leu Met Ser Phe Thr Ala Ala Val Thr Ser Pro			
	1705	1710		1715
	Leu Thr Thr Gln Gln Thr Leu Leu Phe Asn Ile Leu Gly Gly Trp			
	1720	1725		1730

55	56
Val Ala Ala Gln Leu Ala Ala Pro Ala Ala Ala Thr Ala Phe Val	
1735	1740 1745
Gly Ala Gly Ile Thr Gly Ala Val Ile Gly Ser Val Gly Leu Gly	
1750	1755 1760
Lys Val Leu Val Asp Ile Leu Ala Gly Tyr Gly Ala Gly Val Ala	
1765	1770 1775
Gly Ala Leu Val Ala Phe Lys Ile Met Ser Gly Glu Ala Pro Thr	
1780	1715 1790
Ala Glu Asp Leu Val Asn Leu Leu Pro Ala Ile Leu Ser Pro Gly	
1795	1800 1805
Ala Leu Val Val Gly Val Val Cys Ala Ala Ile Leu Arg Arg His	
1810	1815 1820
Val Gly Pro Gly Glu Gly Ala Val Gln Trp Met Asn Arg Leu Ile	
1825	1830 1835
Ala Phe Ala Ser Arg Gly Asn His Val Ser Pro Thr His Tyr Val	
1840	1845 1850
Pro Glu Ser Asp Ala Ser Val Arg Val Thr His Ile Leu Thr Ser	
1855	1860 1865
Leu Thr Val Thr Gln Leu Leu Lys Arg Leu His Val Trp Ile Ser	
1870	1875 1880
Ser Asp Cys Thr Ala Pro Cys Ala Gly Ser Trp Leu Lys Asp Val	
1885	1890 1895
Trp Asp Trp Ile Cys Glu Val Leu Ser Asp Phe Lys Ser Trp Leu	
1900	1905 1910
Lys Ala Lys Leu Met Pro Gln Leu Pro Gly Ile Pro Phe Val Ser	
1915	1920 1925
Cys Gln Arg Gly Tyr Arg Gly Val Trp Arg Gly Glu Gly Ile Met	
1930	1935 1940
His Ala Arg Cys Pro Cys Gly Ala Asp Ile Thr Gly His Val Lys	
1945	1950 1955
Asn Gly Ser Met Arg Ile Val Gly Pro Lys Thr Cys Ser Asn Thr	

57	1960	1965	1970
	Trp Arg Gly Ser Phe Pro Ile Asn Ala His Thr Thr Gly Pro Cys		
	1975	1980	1985
	Thr Pro Ser Pro Ala Pro Asn Tyr Thr Phe Ala Leu Trp Arg Val		
	1990	1995	2000
	Ser Ala Glu Glu Tyr Val Glu Val Arg Arg Leu Gly Asp Phe His		
	2005	2010	2015
	Tyr Ile Thr Gly Val Thr Thr Asp Lys Ile Lys Cys Pro Cys Gln		
	2020	2025	2030
	Val Pro Ser Pro Glu Phe Phe Thr Glu Val Asp Gly Val Arg Leu		
	2035	2040	2045
	His Arg Tyr Ala Pro Pro Cys Lys Pro Leu Leu Arg Asp Glu Val		
	2050	2055	2360
	Thr Phe Ser Ile Gly Leu Asn Glu Tyr Leu Val Gly Ser Gln Leu		
	2065	2070	2375
	Pro Cys Glu Pro Glu Pro Asp Val Ala Val Leu Thr Ser Met Leu		
	2080	2085	2390
	Thr Asp Pro Ser His Ile Thr Ala Glu Thr Ala Ala Arg ArgLeu		
	2005	2040	2445
	Asn Arg Gly Ser Pro Pro Ser Leu Ala Ser Ser Ser Ala Ser Gln		
	2110	2115	2420
	Leu Ser Ala Pro Ser Leu Lys Ala Thr Cys Thr Thr His His Asp		
	2125	2130	2435
	Ser Pro Asp Ala Asp Leu Ile Thr Ala Asn Leu Leu Trp Arg Gln		
	2140	2145	2450
	Glu Met Gly Gly Asn Ile Thr Arg Val Glu Ser Glu Asn Lys Ile		
	2155	2160	2465
	Val Ile Leu Asp Ser Phe Asp Pro Leu Val Ala Glu Glu Asp Asp		
	2170	2175	2480
	Arg Glu Ile Ser Val Pro Ala Glu Ile Leu Leu Lys Ser Lys Lys		
	2185	2190	2495

59

60

The Pro Pro Ala Met Pro Ile Trp Ala Arg Pro Asp Tyr Asn Pro  
 Pro Leu Val Glu Pro Trp Lys Arg Pro Asp Tyr Glu Pro Pro Leu  
 Val His Gly Cys Pro Leu Pro Pro Pro Lys Pro Thr Pro Val Pro  
 Pro Pro Arg Arg Lys Arg Thr Val Val Leu Asp Glu Ser Thr Val  
 Ser Ser Ala Leu Ala Glu Leu Ala Thr Lys Thr Phe Gly Ser Ser  
 Thr Thr Ser Gly Val Thr Ser Gly Glu Ala Ala Glu Ser Ser Pro  
 Ala Pro Ser Cys Asp Gly Glu Leu Asp SerGlu Ala Glu Ser Tyr  
 Ser Ser Met Pro Pro Leu Glu Gly Glu Pro Gly Asp Pro Asp Leu  
 Ser Asp Gly Ser Trp Ser Thr Val Ser Ser Asp Gly Gly Thr Glu  
 Asp Val Val Cys Cys Ser Met Ser Tyr Ser Trp Thr Gly Ala Leu  
 Ile Thr Pro Cys Ala AlaG lu Glu Thr LysLeu Pro Ile Asn Ala  
 Leu Ser Asn Ser Leu Leu Arg His His Asn Leu Val Tyr Ser Thr  
 Thr Ser Arg Ser Ala Gly Gln Arg Gln Lys Lys Val Thr Phe Asp  
 Arg Leu Gln Val Leu Asp Asp His Tyr Arg Asp Val Leu Lys Glu  
 Ala Lys Ala Lys Ala Ser Thr Val Lys AlaLys Leu Leu Ser Val  
 Glu Glu Ala Cys Ser Leu Thr Pro Pro His Ser Ala Arg Ser Lys



61

62

Phe Gly Tyr Gly Ala Lys Asp Val Arg Ser His Ser Ser Lys Ala  
 Glu Glu Ala Cys Ser Leu Thr Pro Pro His Ser Ala Arg Ser Lys  
 Tyr Asp Val Val Lys Gln Leu Pro Ile Ala Val Met Gly Thr Ser  
 Cys Val Lys Pro Glu Lys Gly Gly Arg Lys Pro Ala Arg Leu Ile  
 Val Tyr Pro Asp Leu Gly Val Arg Val Cys Glu Lys Arg Ala Leu  
 Tyr Asp Val Val Lys Gln Leu Pro Ile Ala Val Met Gly Thr Ser  
 Tyr Gly Phe Gln Tyr Ser Pro Ala Gln Arg Val Asp Phe Leu Leu  
 Asn Ala Trp Lys Ser Lys Lys Asn Pro Met Gly Phe Ser Tyr Asp  
 Thr Arg Cys Phe Asp Ser Thr Val Thr Glu Ala Asp Ile Arg Thr  
 Glu Glu Asp Leu Tyr Gln Ser Cys Asp Leu Val Pro Glu Ala Arg  
 Ala Ala Ile Arg Ser Leu Thr Glu Arg Leu Tyr Ile Gly Gly Pro  
 Leu Thr Asn Ser Lys Gly Gln Asn Cys Gly Tyr Arg Arg Cys Arg  
 Ala Ser Gly Val Leu Thr Thr Ser Cys Gly Asn Thr Ile Thr Cys  
 Tyr Leu Lys Ala Ser Ala Ala Cys Arg Ala Ala Lys Leu Arg Asp  
 Cys Thr Met Leu Val Cys Gly Asp Asp Leu Val Val Ile Cys Glu

63

64

Ser Ala Gly Val Gln Glu Asp Ala Ala Asn Leu Arg Ala Phe Thr  
 Glu Ala Met Thr Arg Tyr Ser Ala Pro Pro Gly Asp Pro Pro Gln  
 Pro Glu Tyr Asp Leu Glu Leu Ile Thr Ser Cys Ser Ser Asn Val  
 Ser Val Ala His Asp Gly Ala Gly Lys Arg Val Tyr Tyr Leu Thr  
 Arg Asp Pro Glu Thr Pro Leu Ala Arg Ala Ala Trp Glu Thr Ala  
 Arg His Thr Pro Val Asn Ser Trp Leu Gly Asn Ile Ile Met Phe  
 Ala Pro Thr Leu Trp Val Arg Met Val Leu Met Thr His Phe Phe  
 Ser Ile Leu Ile Ala Gln Glu His Leu Glu Lys Ala Leu Asp Cys  
 Glu Ile Tyr Gly Ala Val His Ser Val Gln Pro Leu Asp Leu Pro  
 Glu Ile Ile Gln Arg Leu His Gly Leu Ser Ala Phe Ser Leu His  
 Ser Tyr Ser Pro Gly Glu Ile Asn Arg Val Ala Ala Cys Leu Arg  
 Lys Leu Gly Val Pro Pro Leu Arg Ala Trp Arg His Arg Ala Arg  
 Ser Val Arg Ala Thr Leu Leu Ser Gln Gly Gly Arg Ala Ala Ile  
 s Gly Lys Tyr Leu Phe Asn Trp Ala Val Lys Thr Lys Leu Lys  
 Leu Thr Pro Leu Pro Ser Ala Ser Gln Leu Asp Leu Ser Asn Trp  
 Phe Thr Gly Gly Tyr Ser Gly Gly Asp Ile Tyr His Ser Val Ser  
 His Val Arg Pro Arg Trp Phe Phe Trp Cys Leu Leu Leu Leu Ser  
 Val Gly Val Gly Ile Tyr Leu Leu Pro Asn Arg

[0047]

配列番号: 16

配列の長さ: 502

配列の型: アミノ酸

配列の種類: 蛋白質 (YS117-5' amino acid)

MetSerThrAsnPro LysProGlnArgLys ThrLysArgAsnThr AsnArgArgProGln  
 AspValLysPhePro GlyGlyGlyGlnIle ValGlyGlyValTyr LeuLeuProArgArg  
 GlyProArgValGly ValArgAlaThrArg LysThrSerGluArg SerGlnProArgGly  
 ArgArgGlnProIle ProLysAlaArgArg ProGluGlyArgSer TrpAlaGlnProGly  
 TyrProTrpProLeu TyrGlyAsnGluGly CysGlyTrpAlaGly TrpLeuLeuSerPro  
 ArgGlySerArgPro SerTrpGlyProThr AspProArgArgArg SerArgAsnLeuGly  
 LysValIleAspThr LeuThrCysGlyPhe AlaAspLeuMetGly TyrIleProLeuVal  
 GlyAlaProLeuGly GlyAlaAlaArgAla LeuAlaHisGlyVal ArgValLeuGluAsp  
 GlyValAsnTyrAla ThrGlyAsnLeuPro GlyCysSerPheSer IlePheLeuLeuAla  
 LeuLeuSerCysLeu ThrValProAlaSer AlaValGluValArg AsnSerSerGlyVal  
 TyrHisValThrAsn AspCysProAsnAla SerValValTyrGlu ThrGluSerLeuIle  
 MetHisLeuProGly CysValProCysVal ArgGluGlyAsnAla SerArgCysTrpVal  
 SerLeuSerProThr IleAlaAlaLysAsp ProSerValProVal SerGluIleArgArg  
 HisValAspLeuIle ValGlyAlaAlaAla PheCysSerAlaMet TyrValGlyAspLeu  
 CysGlySerIlePhe LeuValGlyGlnIle PheThrPheSerPro ArgArgHisTrpThr  
 ThrGlnAspCysAsn CysSerIleTyrPro GlyHisValThrGly HisArgMetAlaTrp  
 AspMetMetMetAsn TrpSerProThrGly AlaLeuValMetAla GlnLeuLeuArgIle  
 ProGlnAlaValVal AspMetIleAlaGly AlaHisTrpGlyVal LeuAlaGlyLeuAla  
 TyrTyrSerMetVal GlyAsnTrpAlaLys ValValValValLeu LeuLeuPheAlaGly  
 ValAspAlaAspThr GlnValThrGlyGly SerAlaAlaTyrAsp AlaArgGlyLeuAla  
 SerLeuPheThrPro GlyProLysGlnAsn IleGlnLeuIleAsn ThrAsnGlySerTrp  
 HisIleAsnArgThr AlaLeuAsnCysAsn GluSerLeuAsnThr GlyTrpValAlaGly  
 LeuPheTyrTyrHis LysPheAsnSerSer GlyCysProGluArg MetAlaSerCysGln  
 ProLeuThrAlaPhe AspGlnGlyTrpGly ProIleThrTyrGlu GlyAsnAlaSerGly  
 AspGlnArgProTyr CysTrpHisTyrAla ProArgProCysGly IleValProAlaArg

GluVal

[0048]

配列番号：17

配列の長さ：365

配列の型：アミノ酸

配列の種類：蛋白質 (YS117-3' amino acid)

ThrValThrGluAla AspIleArgThrGlu GluAspLeuTyrGln SerCysAspLeuVal  
 ProGluAlaArgThr AlaIleArgSerLeu ThrGluArgLeuTyr IleGlyGlyProLeu  
 ThrAsnSerLysGly GlnAsnCysGlyTyr ArgArgCysArgAla SerGlyValLeuThr  
 ThrSerCysGlyAsn ThrIleThrCysTyr LeuLysAlaSerAla AlaCysArgAlaAla  
 LysLeuGlnAspCys ThrMetLeuValCys GlyAspAspLeuVal ValIleCysGluSer  
 AlaGlyValGlnGlu AspAlaAlaSerLeu ArgAlaPheThrGlu AlaMetThrArgTyr  
 SerAlaProProGly AspProProGlnPro GluTyrAspLeuGlu LeuIleThrSerCys  
 SerSerAsnValSer ValAlaHisAspGly SerGlyLysArgVal TyrTyrLeuThrArg  
 AspProGluThrPro LeuAlaArgAlaAla TrpGluThrAlaArg HisThrProValAsn  
 SerTrpLeuGlyAsn IleIleMetPheAla ProThrLeuTrpVal ArgHetValLeuMet  
 ThrHisPhePheSer IleLeuIleAlaGln GluHisLeuGluLys AlaLeuAspCysGlu  
 IleTyrGlyAlaVal HisSerValGlnPro LeuAspLeuProGlu IleIleGlnArgLeu  
 HisGlyLeuSerAla PheSerLeuHisSer TyrSerProGlyGlu IleAsnArgValAla  
 AlaCysLeuArgLys LeuGlyValProPro LeuArgAlaTrpArg HisArgAlaArgSer  
 ValArgAlaThrLeu LeuSerGlnGlyGly LysAlaAlaIleCys GlyLysTyrLeuPhe  
 AsnTrpAlaValLys ThrLysLeuLysLeu ThrProLeuProSer AlaSerGlnLeuAsp  
 LeuSerAsnTrpPhe ThrGlyGlyTyrSer GlyGlyAspIleTyr HisSerValSerHis  
 ValArgProArgTrp PhePheTrpCysLeu LeuLeuLeuSerVal GlyValGlyIleTyr  
 LeuLeuProAsnArg

[0049]

配列番号: 18

配列の長さ: 502

配列の型: アミノ酸

配列の種類: 蛋白質 (SR037-5' amino acid)

MetSerThrAsnPro LysProGlnArgLys ThrLysArgAsnThr AsnArgArgProGln  
 AspValLysPhePro GlyGlyGlyGlnIle ValGlyGlyValTyr LeuLeuProArgArg  
 GlyProArgMetGly ValArgAlaThrArg LysThrSerGluArg SerGlnProArgGly  
 ArgArgGlnProIle ProLysAlaArgArg ProGluGlyArgSer TrpAlaGlnProGly  
 TyrProTrpProLeu TyrGlyAsnGluGly CysGlyTrpAlaGly TrpLeuLeuSerPro  
 ArgGlySerArgPro SerTrpGlyProSer AspProArgArgArg SerArgAsnLeuGly  
 LysValIleAspThr LeuThrCysGlyPhe AlaAspLeuMetGly TyrIleProLeuVal  
 GlyAlaProLeuGly GlyAlaAlaArgAla LeuAlaHisGlyVal ArgValLeuGluAsp  
 GlyValAsnTyrAla ThrGlyAsnLeuPro GlyCysSerPheSer IlePheLeuLeuAla  
 LeuLeuSerCysLeu ThrValProAlaSer AlaValGlyValArg AsnSerSerGlyVal  
 TyrHisValThrAsn AspCysProAsnAla SerValValTyrGlu ThrGluSerLeuIle  
 IleHisLeuProGly CysValProCysVal ArgGluGlyAsnAla SerArgCysTrpVal  
 SerLeuSerProThr ValAlaAlaLysAsp ProSerValProVal SerGluIleArgArg  
 HisValAspLeuIle ValGlyAlaAlaAla PheCysSerAlaMet TyrValGlyAspLeu  
 CysGlySerIlePhe LeuValGlyGlnIle PheThrLeuSerPro ArgArgHisTrpThr  
 ThrGlnAspCysAsn CysSerIleTyrPro GlyHisValThrGly HisArgMetAlaTrp  
 AspMetMetMetAsn TrpSerProThrGly AlaLeuValValAla GlnLeuLeuArgIle  
 ProGlnAlaValVal AspMetIleAlaGly AlaHisTrpGlyVal LeuAlaGlyLeuAla  
 TyrTyrSerMetVal GlyAsnTrpAlaLys ValValValValLeu LeuLeuPheAlaGly  
 ValAspAlaGluThr GlnValSerGlyGly SerAlaAlaGlnThr ThrTyrGlyLeuThr  
 AlaLeuPheArgThr GlyProAsnGlnLys IleGlnLeuIleAsn ThrAsnGlySerTrp  
 HisIleAsnArgThr AlaLeuAsnCysAsn GluSerLeuHisThr GlyTrpLeuAlaAla  
 LeuPheTyrThrHis LysPheAsnSerSer GlyCysLeuGluArg MetAlaSerCysGln  
 ProLeuSerAlaPhe AspGlnGlyTrpGly ProIleThrTyrGly GlyAsnAlaSerAsp  
 GlyGlnArgProTyr CysTrpHisTyrAla ProArgProCysGly IleValProAlaArg

GluVal

[0050]

配列番号：19

配列の長さ：365

配列の型：アミノ酸

配列の種類：蛋白質 (SR037-3' amino acid)

ThrValThrGluAla AspIleArgThrGlu GluAspLeuTyrGln SerCysAspLeuVal  
 ProGluAlaArgThr AlaIleArgSerLeu ThrGluArgLeuTyr IleGlyGlyProLeu  
 ThrAsnSerLysGly GlnAsnCysGlyTyr ArgArgCysArgAla SerGlyValLeuThr  
 ThrSerCysGlyAsn ThrIleThrCysTyr LeuLysAlaSerAla AlaCysArgAlaAla  
 LysLeuArgAspCys ThrMetLeuValCys GlyAspAspLeuVal ValIleCysGluSer  
 AlaGlyValGlnGlu AspAlaAlaSerLeu ArgAlaPheThrGlu AlaMetThrArgTyr  
 SerAlaProProGly AspProProGlnPro GluTyrAspLeuGlu LeuIleThrSerCys  
 SerSerAsnValSer ValAlaHisAspGly AlaGlyLysArgVal TyrTyrLeuThrArg  
 AspProGluThrPro LeuAlaArgAlaAla TrpGluThrAlaArg HisThrProValAsn  
 SerTrpLeuGlyAsn IleIleMetPheAla ProThrLeuTrpVal ArgMetValLeuMet  
 ThrHisPhePheSer IleLeuIleAlaGln GluHisLeuGluLys AlaLeuAspCysGlu  
 IleTyrGlyAlaVal HisSerIleGlnPro LeuAspLeuProGlu IleIleGlnArgLeu  
 HisGlyLeuSerAla PheSerLeuHisSer TyrSerProGlyGlu IleAsnArgValAla  
 AlaCysLeuArgLys LeuGlyValProPro LeuArgAlaTrpArg HisArgAlaArgSer  
 ValArgAlaThrLeu LeuSerGlnGlyGly LysAlaAlaIleCys GlyLysTyrLeuPhe  
 AsnTrpAlaValLys ThrLysLeuLysLeu ThrProLeuProSer AlaSerGlnLeuAsp  
 LeuSerAsnTrpPhe ThrGlyGlyTyrAsn GlyGlyAspIleTyr HisSerValSerArg  
 ValArgProArgTrp PhePheTrpCysLeu LeuLeuLeuSerVal GlyValGlyIleTyr  
 LeuLeuProAsnArg

【表1】

## 配列決定に使用したプライマーの配列ならびに位置

番号	配列 (5' 端より3' 端へ)	5' 端からの位置
#345	5'-GTTGGCACTATGCCCTACGC-3'	nt1798-1817 (S)
#344	GCAGGAGTTTGGTGATGTCA	nt2999-3018 (A)
#347	TGTGTGCTCCACCGCACIT	nt2973-2992 (S)
#352	TATCGCCCGACGCGCACT	nt3864-3883 (A)
#350	TCGGACCTTTACTTGGTCAC	nt3723-3742 (S)
#349	GGCTGAAATCGACTGTCTGG	nt4718-4737 (A)
#354	GTTTTGGGAGAGCGTCTTCA	nt5009-5028 (S)
#356	AGCGACTGACGCGTCGCTCT	nt6151-6170 (A)
#342	TCCCCACGCACTACGTGCC	nt6129-6148 (S)
#344	GTGATGTTGCCGCCCATCTC	nt7049-7066 (A)
#358	ACAGCCAACCTCCTGTGGAG	nt7023-7042 (S)
#360	CTGTGGATGCCTTGGCCTTA	nt7814-7833 (A)
#335	ACCAGGACGTGCTCAAGGAG	nt7792-7811 (S)
#316	CATAGGCTCCGTGAAGGCTC	nt8611-8630 (A)

(S) : センス、(A) : アンチセンス

【表2】

## HC-G9、YS117、SR037間の領域別塩基配列相同性

領域	塩基長 (アミノ酸長)	配列の相同性 (%)	
		塩基配列	アミノ酸配列
5' 非翻訳域	259	99.6-100	
コア領域	573 (191)	96.2-97.4	99.0-99.5
エンベロープ	576 (192)	92.9-95.0	93.2-97.4
E2/NS1	357 (119)	84.3-87.4	76.5-82.4
NS5	1096 (365)	96.5-97.6	98.6-98.9
3' 非翻訳域	66	90.5-92.1	
合 計	2927 (867)	94.5-95.6	94.7-96.2

【表3】

## 1c型HCVの出現頻度

国名	1c型HCVの出現率
	1c型の数/検査総数 (%)
日本	0/670 (0%)
中華人民共和国	0/262 (0%)
タイ	0/64 (0%)
インドネシア	8/81 (9.9%)
ニュージーランド	0/60 (0%)

【図面の簡単な説明】

50 図1: HCV遺伝子配列決定に利用した増幅領域を示す

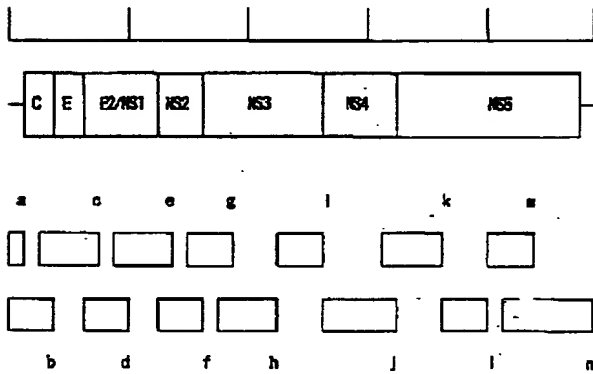
図

図2: HC-G9とこれまでに全域の塩基配列が解明されている14のHCV株との塩基配列の相同性を示す図。

図3: 実施例2にて行った電気泳動試験の泳動パターンを示す電気泳動写真。(a)に従来法による判定を示

【図1】

塩基配列決定に利用したHCV領域



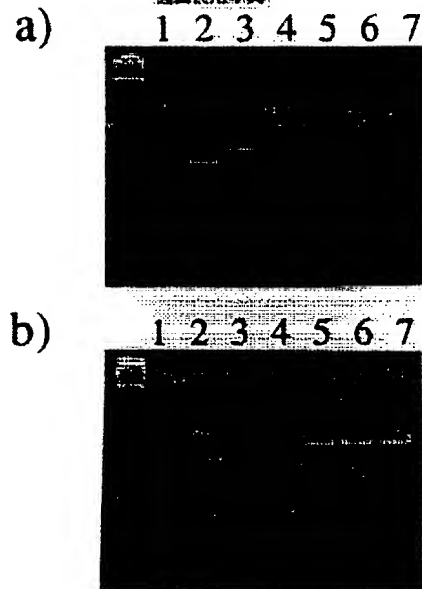
a (nt1-160) ; b (nt63-847) ; c (nt732-1606) ; d (nt1300-1857) ; e (nt1798-2550)  
 f (nt241-3018) ; g (nt2973-3383) ; h (nt3723-4737) ; i (nt4883-5050)  
 j (nt5009-6170) ; k (nt6129-7088) ; l (nt7023-7833) ; m (nt7732-8630)  
 n (nt8259-9440)

左端に5' 端、右に3' 端が来るように示している。  
 塩基配列は5' 端からの塩基数 (nt) で示した。

す。1 c 型は検出されない。(b) に #321 と #104 での判定を示す。1 c 型以外は検出されない。1 は HC-J 1、2 は HC-J 4、3 は HC-J 6、4 は HC-J 8、5 は HC-G 9、6 は YS 117、7 は SR 037 を示す。

【図3】

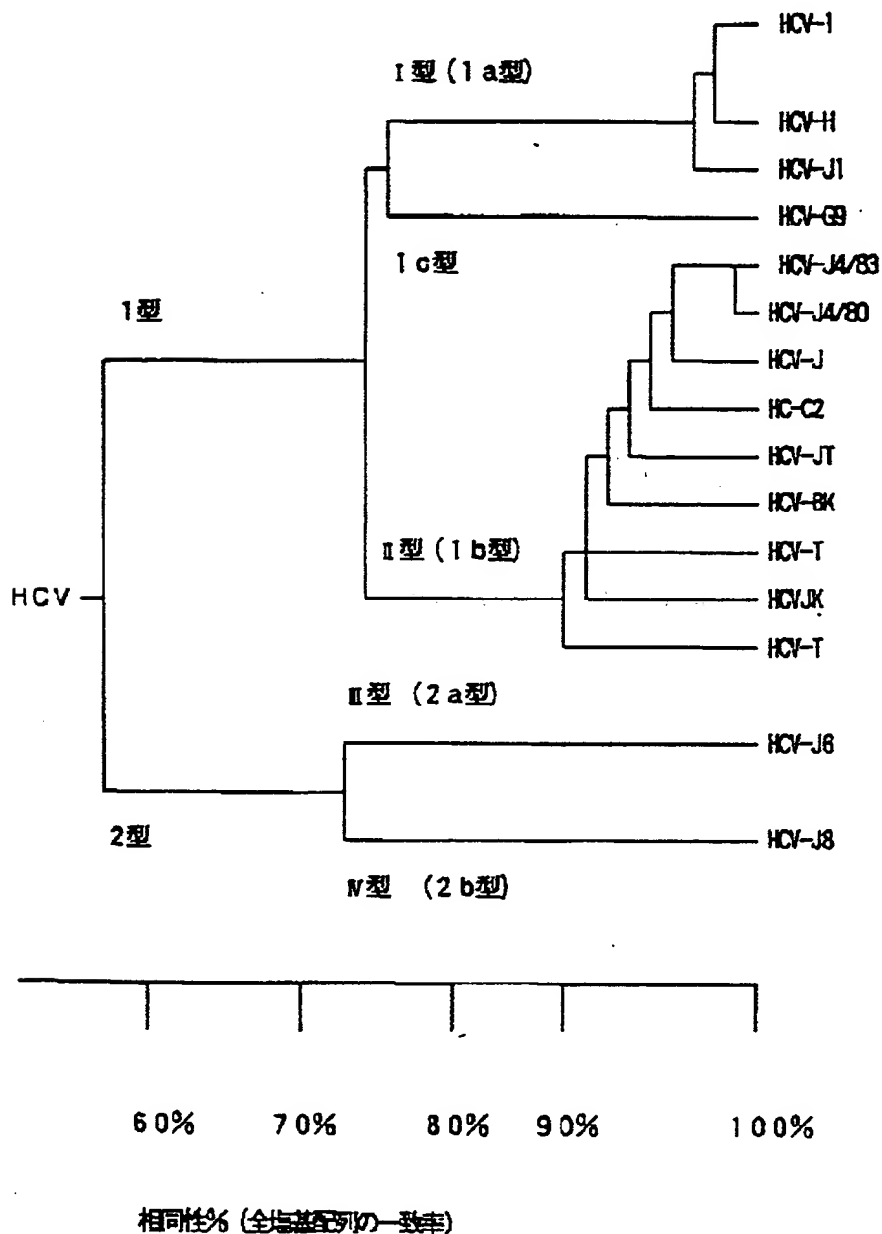
各遺伝子型サンプルの電気泳動写真





【図2】

これまでに全塩基配列が公表されたHCV株との塩基配列の相  
性



フロントページの続き

(51) Int. Cl. <sup>5</sup>

C12Q 1/68

G01N 33/576

識別記号

庁内整理番号

F I

技術表示箇所

A 7823-4B

Z 8310-2J

(41)

特開平6-319563

// G 0 1 N 33/53  
C 0 7 K 99:00

D 8310-2 J